Reachability, observability, minimality and extendibility for two-point boundary-value descriptor systems

Ramine Nikoukhah, A.S. Willsky, Bernard C. Levy

UNITÉ DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél. (1) 39 63 55 11

Rapports de Recherche

N° 920

REACHABILITY, OBSERVABILITY, MINIMALITY AND EXTENDIBILITY FOR TWO-POINT BOUNDARY-VALUE DESCRIPTOR SYSTEMS

Programme 5

Ramine NIKOUKHAH
Alan S WILLSKY
Bernard C LEVY

Octobre 1988

REACHABILITY, OBSERVABILITY, MINIMALITY AND EXTENDIBILITY FOR TWO-POINT BOUNDARY-VALUE DESCRIPTOR SYSTEMS

Ramine Nikoukhah
INRIA, Rocquencourt
78153 Le Chesnay Cedex
France

Alan S. Willsky
LIDS, M.I.T.
Cambridge, MA 02139
USA

Bernard C. Levy
University of California
Davis, CA 95616
USA

Abstract:
A deterministic system theory is developed for two-point boundary-value descriptor systems (TPBVDS's). In particular, detailed characterizations of the properties of reachability, observability and minimaltiy are obtained. In addition, extendibility, i.e. the concept of considering a TPBVDS as being defined on a sequence of intervals of increasing length, is defined and studied. These system-theoretic properties are derived for general TPBVDS's and then specialized to the case of stationary systems for which the input-output map (weighting pattern) is shift-invariant.

UNE THEORIE DE LA COMMANDABILITE, L'OBSERVABILITE, LA MINIMALITE ET L'EXTENSIBILITE POUR LES SYSTEMES DYNAMIQUES AUX DEUX BOUTS

Résumé:
Une théorie déterministe est développée pour les systèmes linéaires dynamique aux deux bouts. En particulier, les notions de commandabilité, d'observabilité et de minimalité sont étudiées en détail. La notion d'extensibilité, consistant à étudier le système sur des intervalles de tailles croissantes, est introduite. Les résultats sont obtenus dans le cas général et ils sont ensuite spécialisés au cas stationnaire — cas où la relation entrée-sortie est invariante en temps.

1–Introduction

In this report we study the system–theoretic properties of two–point boundary–value descriptor systems (TPBVDS's) and two related classes of shift–invariant two–point boundary–value descriptor systems namely displacement systems for which the Green's function is shift–invariant, and stationary systems for which the input–output map is stationary. We present detailed characterizations of the properties of strong and weak reachability and observability introduced in [1] and of minimality as well. Another property that is studied in this report is that of extendibility, i.e. the concept of considering a TPBVDS as being defined on a sequence of intervals of increasing length.

Some of the results in this report, such as results concerning the concepts of well–posedness, standard–form, generalized Cayley–Hamilton theorem, inward and outward processes, and strong reachability and observability have already been discussed in [1,2] and so we shall simply review them here. Other results, such as those concerning extendibility, weak reachability and observability, and minimality have only been considered in a much more restricted setting (essentially in the displacement case). The attempt to generalize these concepts have generally failed because no closed–form expressions for the inward process could be found. Here we shall obtain the necessary closed–form expressions and completely resolve the problems of extendibility, weak reachability and observability, and minimality in the most general case.

In the next section we introduce TPBVDS's, and define and characterize two notions of shift–invariant systems, namely displacement systems and stationary systems.

In Section 3 we review the notions of inward and outward processes introduced for TPBVDS's in [1,2] and characterize these processes. We also introduce the concept of extendibility and characterize this property. Section 4 discusses the properties of reachability and observability for TPBVDS's, while in Section 5 we present minimality results. Some extensions are presented in Section 6, and we conclude with a brief discussion in Section 7.

2–Two–Point Boundary–Value Descriptor Systems

A TPBVDS is described by the following dynamic equation

$$Ex(k+1) = Ax(k) + Bu(k), \quad 0 \le k \le N-1 \tag{2.1}$$

with boundary condition

$$V_i x(0) + V_f x(N) = v \tag{2.2}$$

and output

$$y(k) = Cx(k), \quad k=0,1,\ldots,N. \tag{2.3}$$

Here x and v are n–dimensional, u is m–dimensional, y is p–dimensional, and E, A, B, V_i, V_f, and C are constant matrices. In [1] it is shown that if (2.1)–(2.2) is well–posed (i.e. it yields a well–defined map from {u,v} to x), we can assume, without loss of generality that (2.1)–(2.2) is in <u>normalized form</u>, i.e. that there exist scalars α and β such that

$$\alpha E+\beta A = I \tag{2.4}$$

(this is referred to as the <u>standard form</u> for the pencil {E,A}) and in addition

$$V_i E^N + V_f A^N = I . \tag{2.5}$$

Note that (2.4) implies that E and A commute, that E, A and the system have a common set of eigenvectors[1], and also that $\{E^k,A^k\}$ is a regular pencil for all $k \ge 0$ (see[1]). But most importantly (2.4) implies that the space of matrices $A^K E^L$, $K,L \ge 0$, is spanned by the n matrices $\{A^k E^{n-1-k} \mid k=0,..,n-1\}$; this property has been introduced in [1] as the generalized Cayley–Hamilton theorem. We assume throughout this paper that (2.4) and (2.5) hold. We also assume that the interval of definition of our system is sufficiently large to excite and observe all system modes. Specifically, we assume that $N \ge 2n$, unless explicitly stated otherwise.

As derived in [1], the map from {u,v} to x has the following form:

[1]v is an eigenvector of the system if $v \ne 0$ and for some σ, $(\sigma E-A)v=0$. σ is called an eigenmode of the system; for descriptor systems σ can be ∞ as well.

$$x(k) = A^k E^{N-k} v + \sum_{j=0}^{N-1} G(k,j)Bu(j), \tag{2.6}$$

where the Green's function G(k,j) is given by

$$G(k,j) = \begin{cases} A^k(A-E^{N-k}(V_iA+\omega V_fE)E^k)E^{j-k}A^{N-j-1}\Gamma^{-1} & j \geq k \\ E^{N-k}(\omega E-A^k(V_iA+\omega V_fE)A^{N-k})E^j A^{k-j-1}\Gamma^{-1} & j<k \end{cases} \tag{2.7}$$

and where ω is any number such that

$$\Gamma \triangleq \omega E^{N+1} - A^{N+1} \tag{2.8}$$

is invertible.

In marked contrast to the case for causal systems ($E=I, V_f=0$), G(k,j) does <u>not</u>, in general, depend on the difference of its arguments. Borrowing some terminology from [11–13], we have the following definition of our first notion of shift–invariance:

<u>Definition 2.1</u>

The TPBVDS (2.1)–(2.2) is a <u>displacement system</u> if (with the usual abuse of notation) for $0 \leq k \leq N$, $0 \leq j \leq N-1$

$$G(k,j) = G(k-j). \tag{2.9}$$

A second notion of shift–invariance is the one associated with the input–output map. Specifically, with v=0 in (2.2), we have that (2.1)–(2.3) defines a linear map of the form

$$y(k) = \sum_{j=0}^{N-1} W(k,j)u(j), \tag{2.10}$$

where, obviously

$$W(k,j) = CG(k,j)B. \tag{2.11}$$

Definition 2.2

The TPBVDS (2.1)–(2.3) is stationary if (again with the usual abuse of notation)

$$W(k,j) = W(k-j) \tag{2.12}$$

for $0 \le k \le N$, $0 \le j \le N-1$.

The following results characterize the conditions under which a TPBVDS is displacement and stationary.

Theorem 2.1

The TPBVDS (2.1)–(2.3) is stationary if and only if

$$O_s[V_i,E]R_s = O_s[V_i,A]R_s = 0 \tag{2.13a}$$

$$O_s[V_f,E]R_s = O_s[V_f,A]R_s = 0, \tag{2.13b}$$

where [X,Y] denotes the commutator product of X and Y

$$[X,Y] = XY - YX \tag{2.14}$$

and

$$R_s = [A^{n-1}B|EA^{n-2}B|\ldots|E^{n-1}B] \tag{2.15}$$

$$O_s = \begin{bmatrix} CA^{n-1} \\ CEA^{n-2} \\ \vdots \\ CE^{n-1} \end{bmatrix}. \tag{2.16}$$

Before proving this result, let us also state a corollary (which will also require proof) and make several comments:

Corollary

The TPBVDS (2.1)–(2.2) is a displacement system if and only if

$$[V_i,E] = [V_i,A] = 0 \tag{2.17a}$$

$$[V_f,E] = [V_f,A] = 0. \tag{2.17b}$$

The matrices R_s and O_s in (2.15), (2.16) are, respectively, the strong reachability and strong observability matrices of the TPBVDS as discussed in [1] (see also Section 4). Thus (2.13) states that V_i and V_f must commute with E and A except for parts that are either in the left nullspace of R_s or the right nullspace of O_s. For example, if R_s and O_s are of full rank – i.e. if the TPBVDS is strongly reachable and strongly observable – V_i and V_f must commute with E and A. Turning to the corollary, we see that this is precisely the condition for a TPBVDS to be displacement. Thus as expected from (2.11), a displacement system is always stationary. Furthermore, the only way in which a TPBVDS can be stationary without being a displacement system is if the system is not strongly reachable or strongly observable.

The results of causal system theory might then suggest that this distinction between displacement and stationary is a trivial artifact caused by the use of possible non–minimal realizations. However, as in the case of continuous–time boundary–value systems [7], we will see that the story is different for TPBVDS. Specifically, as will be shown in Section 5, a TPBVDS can be minimal without being strongly reachable or strongly observable.

Proof of the Corollary

Assume that Theorem 2.1 holds. Then, from (2.11) we see that the concepts of stationarity and displacement are the same if C=B=I. Thus from Theorem 2.1, a TPBVDS is displacement if and only if (2.13) holds with R_s and O_s defined with C=B=I. However, thanks to the generalized Cayley–Hamilton theorem for pencils in standard form [1], the matrices $\{A^kE^{n-k-1} \mid k=0,\ldots n-1\}$ span the same set as $\{E^kA^j \mid k,j\geq 0\}$. Thus R_s and O_s are of full rank, so that (2.13) is equivalent to (2.17).

Proof of Theorem 2.1

What we must show is that (2.13) is equivalent to

$$W(k+1,j+1) = W(k,j) \tag{2.18}$$

for $0 \le k \le N-1$, $0 \le j \le N-2$. Then, using (2.7), the commutativity of E and A, and performing some algebra we find that (2.18) is equivalent to

$$\begin{aligned} &CA^{k+1}E^{N-k-1}[V_iA+\omega V_fE]A^{N-j-2}E^{j+1}\Gamma^{-1}B = \\ &CA^kE^{N-k}[V_iA+\omega V_fE]A^{N-j-1}E^j\Gamma^{-1}B. \end{aligned} \tag{2.19}$$

Now thanks to the generalized Cayley–Hamilton theorem and to (2.4), we have that the set of matrices

$$\{A^kE^{M-k} \mid k=0,1,\dots,M\}$$

spans the same space as

$$\{A^kE^{n-k-1} \mid k=0,1,\dots,n-1\}$$

as long as $M \ge n-1$. Thus, since $N \ge 2n$, (2.19) yields

$$O_sA[V_iA+\omega V_fE]E\Gamma^{-1}R_s = O_sE[V_iA+\omega V_fE]A\Gamma^{-1}R_s. \tag{2.20}$$

The range of the matrix $\Gamma^{-1}R_s$ is independent of ω. To see this, define the strong reachability subspace

$$\mathcal{R}_s = \mathrm{Im}(R_s). \tag{2.21}$$

Then the generalized Cayley–Hamilton theorem implies that for all $M \ge 0$

$$A^M\mathcal{R}_s \subset \mathcal{R}_s \quad , \quad E^M\mathcal{R}_s \subset \mathcal{R}_s \tag{2.22}$$

which implies

$$(\omega E^{N+1} - A^{N+1})\mathcal{R}_s = \Gamma\mathcal{R}_s \subset \mathcal{R}_s. \tag{2.23}$$

Thus $\mathcal{R}_s$ is Γ–invariant which implies that $\mathcal{R}_s$ is Γ^{-1}–invariant. In fact, for all ω such that Γ^{-1} exists

$$\Gamma^{-1}\mathcal{R}_s = \Gamma\mathcal{R}_s = \mathcal{R}_s. \tag{2.24}$$

Since the range of $\Gamma^{-1}R_s$ does not depend on ω and (2.20) must hold for almost all values of ω (i.e. all values for which Γ is invertible), we can deduce that (2.20) is equivalent to the following pair of equalities

$$O_s[AV_iE-EV_iA]A\Gamma^{-1}R_s = 0 \tag{2.25}$$

$$O_s[AV_fE-EV_fA]E\Gamma^{-1}R_s = 0. \tag{2.26}$$

Now, note that (2.25) is equivalent to the pair of equalities

$$O_s[AV_iE-EV_iA]A\Gamma^{-1}A^NR_s = 0 \tag{2.27}$$

$$O_s[AV_iE-EV_iA]A\Gamma^{-1}E^NR_s = 0. \tag{2.28}$$

To see this observe that since $\{E^N,A^N\}$ is regular

$$\mathcal{R}_s = \mathrm{Im}([A^NR_s \mid E^NR_s]). \tag{2.29}$$

In a similar fashion we have that (2.26) is equivalent to the pair of equalities

$$O_s[AV_fE-EV_fA]E\Gamma^{-1}A^NR_s = 0 \tag{2.30}$$

$$O_s[AV_fE-EV_fA]E\Gamma^{-1}E^NR_s = 0. \tag{2.31}$$

Using the commutativity of E and A together with (2.5), we see that (2.30) is equivalent to

$$O_s[-AV_iE+EV_iA]E^{N+1}\Gamma^{-1}R_s = 0. \tag{2.32}$$

Using the definition of Γ, we see that (2.27) and (2.32) imply that

$$O_s[AV_iE-EV_iA]R_s = 0. \tag{2.33}$$

In a similar fashion (2.28) and (2.31) can be shown to imply

$$O_s[AV_fE-EV_fA]R_s = 0. \tag{2.34}$$

Note also that

$$E\Gamma^{-1}\mathcal{R}_s \subset \mathcal{R}_s \ , \ A\Gamma^{-1}\mathcal{R}_s \subset \mathcal{R}_s \tag{2.35}$$

so that (2.33), (2.34) imply and thus are equivalent to (2.25), (2.26).

Finally, note that thanks to the commutativity of E and A, (2.13a) implies (2.33) and (2.13b) implies (2.34). To see that the reverse of these implications holds, suppose that $\alpha\neq 0$ in (2.4) (if $\alpha=0$, reverse the role of E and A in what follows). Then

$$E = \gamma I + \delta A \ , \quad \gamma\neq 0. \tag{2.36}$$

Substituting this into (2.33) yields

$$O_s[V_i,A]R_s = 0, \tag{2.37}$$

and (2.36) implies

$$O_s[V_i,E]R_s = 0. \tag{2.38}$$

Similarly (2.34) implies (2.13b), and the theorem is proved.

As we shall see, the characterization of the displacement property in (2.13) simplifies many of the computations associated with TPBVDS's. In particular, it is not difficult to check that the Green's function of a displacement system is given by

$$G(k) = \begin{cases} V_i A^{k-1} E^{N-k} & k>0 \\ -V_f E^{-k} A^{N+k-1} & k\leq 0 \end{cases} \tag{2.39}$$

Similarly, the weighting pattern of a stationary TPBVDS is given by

$$W(k) = \begin{cases} CV_i A^{k-1} E^{N-k} B & k>0 \\ -CV_f E^{-k} A^{N+k-1} B & k\leq 0 \end{cases} . \tag{2.40}$$

Before closing this section we consider another problem, namely that of the degree of freedom in the choice of boundary matrices V_i and V_f.

Theorem 2.2

Consider two TPBVDS's with the same matrices C, E, A, and B and identical weighting patterns. Then, if one has boundary matrices V_i and V_f, and the other $\hat{V}_i$ and $\hat{V}_f$, we must have

$$O_s V_i R_s = O_s \hat{V}_i R_s \tag{2.41a}$$

$$O_s V_f R_s = O_s \hat{V}_f R_s . \tag{2.41b}$$

Conversly if (2.41) holds for two TPBVDS's with identical C, E, A, and B system matrices, then their weighting patterns must be identical.

Proof

By setting

$$W(k,j) = \hat{W}(k,j) \tag{2.42}$$

we get that

$$O_sV_iA\Gamma^{-1}R_s = O_s\hat{V}_iA\Gamma^{-1}R_s \tag{2.43a}$$

$$O_sV_fE\Gamma^{-1}R_s = O_s\hat{V}_fE\Gamma^{-1}R_s \tag{2.43b}$$

which in turn implies that

$$O_sV_iA^{N+1}\Gamma^{-1}R_s = O_s\hat{V}_iA^{N+1}\Gamma^{-1}R_s \tag{2.44a}$$

$$O_sV_fA^NE\Gamma^{-1}R_s = O_s\hat{V}_fA^NE\Gamma^{-1}R_s. \tag{2.44b}$$

Now using the fact that both systems are in normalized form, we can rewrite (2,44b) as follows:

$$O_sV_iE^{N+1}\Gamma^{-1}R_s = O_s\hat{V}_iE^{N+1}\Gamma^{-1}R_s \tag{2.45}$$

which in conjunction with (2.44a) implies (2.41a). Equation (2.41b) can be verified similarly.

The converse is easy to see to be true by noting that in the expression for W, V_i and V_f appear only premultiplied by O_s and postmultiplied by R_s.

3– Inward Processes, Outward Processes, and Extendibility

As discussed in [1] (with motivation from [7]) the process x in a TPBVDS can be recovered from two processes that each have interpretations as state processes. The outward process, which expands outward toward the boundaries, summarizes all that one needs to know about the input inside any interval in order to determine x outside the interval. The inward process uses input values near the boundary to propagate the boundary condition inward.

The outward process has a simple definition and characterization [1]. In Krener's context the outward process represents the "jump" corresponding to the difference between x at one end of any interval and the value predicted for x at that point given x at the other end of the interval and assuming zero input inside the interval. In our context, we can't predict in either direction (due to the possible singularity of E and A) and thus we use a somewhat different definition

$$z_o(k,j) = E^{j-k}x(j) - A^{j-k}x(k) \qquad k<j. \tag{3.1}$$

This agrees with Krener's definition if E=I but in general in our case $z_o(k,j)$ can only be propagated in an outward direction. Also, it is possible to write a closed–form expression in terms of the intervening inputs:

$$z_o(k,j) = \sum_{r=k}^{j-1} E^{r-k}A^{j-r-1}Bu(r) \tag{3.2}$$

and to write outward recursions

$$z_o(k-1,j) = Ez_o(k,j) + A^{j-k}Bu(k-1) \tag{3.3a}$$

$$z_o(k,j+1) = Az_o(k,j) + E^{j-k}Bu(j)\ . \tag{3.3b}$$

Note that $z_o(k,j)$ does not involve the boundary matrices V_i and V_f – i.e. it only involves (2.1) – and thus none of the expressions (3.1)–(3.3b) are affected if the TPBVDS has the displacement property.

The situation is considerably different, however, for the inward process. As developed in [1], for any K≤L the inward process $z_i(K,L)$ is a function of the boundary value v and the inputs $\{u(0),...,u(K-1)\}$ and $\{u(L),...,u(N-1)\}$ so that the TPBVDS

$$Ex(k+1) = Ax(k) + Bu(k) \tag{3.4a}$$

$$V_i(K,L)x(K) + V_f(K,L)x(L) = z_i(K,L) \tag{3.4b}$$

yields the same solution as (2.1), (2.2) for $K \le k \le L$. Here $V_i(K,L)$ and $V_f(K,L)$ are assumed to be such that (3.4) is in normalized form, i.e.

$$V_i(K,L)E^{L-K} + V_f(K,L)A^{L-K} = I. \tag{3.5}$$

Theorem 3.1

The inwardly–propagated boundary matrices and the inward process can be expressed as follows:

$$V_i(K,L) = E^{N-L}(\omega E - A^K(\omega V_f E + V_i A)A^{N-K})\Gamma^{-1}E^K \tag{3.6a}$$

$$V_f(K,L) = -A^K(A - E^{N-L}(\omega V_f E + V_i A)E^L)\Gamma^{-1}A^{N-L} \tag{3.6b}$$

and

$$\begin{aligned} z_i(K,L) = {} & E^{N-L}A^K v + E^{N-L}(\omega E - A^K(\omega V_f E + V_i A)A^{N-K})\Gamma^{-1}z_o(0,K) \\ & + A^K(A - E^{N-L}(\omega V_f E + V_i A)E^L)\Gamma^{-1}z_o(L,N). \end{aligned} \tag{3.7}$$

Note in particular the starting values

$$z_i(0,N)=v,\ V_i(0,N)=V_i,\ V_f(0,N)=V_f \tag{3.8}$$

and the "final values"

$$z_i(k,k) = x(k) \qquad \text{for all k.} \tag{3.9}$$

Proof

Let S_h be the following hx(h+1) block matrix,

$$S_h = \begin{bmatrix} -A & E & 0 & .. & 0 \\ 0 & -A & E & .. & 0 \\ : & & & & : \\ 0 & 0 & & -A & E \end{bmatrix} \tag{3.10}$$

then (2.1)–(2.2) can be expressed as

$$\begin{bmatrix} & S_N & \\ V_i\ 0 \dots & 0 & V_f \end{bmatrix} \begin{bmatrix} x(0) \\ \vdots \\ x(N) \end{bmatrix} = \begin{bmatrix} Bu(0) \\ \vdots \\ Bu(N\text{-}1) \\ v \end{bmatrix}. \tag{3.11}$$

Also it is easy to see that

$$\begin{bmatrix} -A^K & E^K & 0 & \dots & 0 \\ 0 & \vdots & \cdots & \vdots & \vdots \\ \vdots & \vdots & S_{L\text{-}K} & \vdots & 0 \\ 0 & \dots & 0 & -A^{N\text{-}L} & E^{N-L} \\ V_i & 0 & \dots & 0 & V_f \end{bmatrix} \begin{bmatrix} x(0) \\ x(K) \\ x(K+1) \\ \vdots \\ x(L) \\ x(N) \end{bmatrix} = \begin{bmatrix} z_o(0,K) \\ Bu(K) \\ \vdots \\ Bu(L\text{-}1) \\ z_o(L,N) \\ v \end{bmatrix}. \tag{3.12}$$

Now to find $V_i(K,L)$ and $V_f(K,L)$, we need to first construct a full–rank matrix

$$[T_i(K,L)\ T_f(K,L)\ P(K,L)]$$

so that

$$[T_i(K,L)\ T_f(K,L)\ P(K,L)] \begin{bmatrix} -A^K & 0 \\ 0 & E^{N\text{-}L} \\ V_i & V_f \end{bmatrix} = 0. \tag{3.13}$$

If we now premultiply both sides of (3.12) by

$$\Omega(K,L) = \begin{bmatrix} 0 & I & 0 & 0 \\ T_i(K,L) & 0 & T_f(K,L) & P(K,L) \end{bmatrix} \tag{3.14}$$

we obtain

$$\begin{bmatrix} & S_{L-K} & \\ T_i(K,L)E^K & & -T_f(K,L)A^{N-L} \end{bmatrix} \begin{bmatrix} x(K) \\ \vdots \\ x(L) \end{bmatrix} = \begin{bmatrix} Bu(K) \\ \vdots \\ Bu(L-1) \\ T_i(K,L)z_o(0,K)+T_f(K,L)z_o(L,N)+P(K,L)v \end{bmatrix}. \tag{3.15}$$

then clearly,

$$V_i(K,L) = T_i(K,L)E^K \tag{3.16a}$$

$$V_f(K,L) = -T_f(K,L)A^{N-L} \tag{3.16b}$$

and

$$z_i(K,L) = T_i(K,L)z_o(0,K)+T_f(K,L)z_o(L,N)+P(K,L)v. \tag{3.17}$$

It is straightforward to check that

$$T_i(K,L) = E^{N-L}(\omega E-A^K(\omega V_fE+V_iA)A^{N-K})\Gamma^{-1} \tag{3.18a}$$

$$T_f(K,L) = A^K(A-E^{N-L}(\omega V_f E+V_i A)E^L)\Gamma^{-1} \tag{3.18b}$$

and

$$P(K,L) = E^{N-L}A^K \tag{3.18c}$$

satisfy our requirements. Clearly then (3.6) and (3.7) can be obtained from (3.18).

Theorem 3.1 can be slightly generalized to give us a relationship between all inwardly–propagated boundary matrices:

$$V_i(K,L) = E^{J-L}(\omega E-A^{K-I}(\omega V_f(I,J)E+V_i(I,J)A)A^{J-K})\Gamma_{J-I}^{-1}E^{K-I} \tag{3.19a}$$

$$V_f(K,L) = -A^{K-I}(A-E^{J-L}(\omega V_f(I,J)E+V_i(I,J)A)E^{L-I})\Gamma_{J-I}^{-1}A^{J-L} \tag{3.19b}$$

when [K,L] is contained in [I,J] and where

$$\Gamma_M = \omega E^{M+1} - A^{M+1}.$$

Corollary

Suppose that (2.1)–(2.2) is a displacement system. Then,

$$V_i(K,L) = V_i E^{N-L+K} \tag{3.20a}$$

$$V_f(K,L) = V_f A^{N-L+K} \tag{3.20b}$$

and

$$z_i(K,L) = E^{N-L}A^K v + V_i E^{N-L} z_o(0,K) - V_f A^K z_o(L,N). \tag{3.21}$$

Proof

Equations (3.20) and (3.21) are easily derived from (3.6) and (3.7) using the fact that E and A must commute with V_i and V_f.

An important interpretation of the inward process, or more specifically the inwardly–propagated boundary matrices (3.6) is that the Green's function for the system (3.4) on the smaller interval [K,L] is the restriction of the Green's function of the original system (2.1)–(2.2) defined on [0,N]. A logical question then is whether we can also move the boundary conditions outward so that the Green's function for the resulting system, when restricted to [0,N] yields the original Green's function. This is roughly the property of extendibility. In particular, it makes a good deal of sense to consider extendibility when one considers shift–invariance, as the intuitive notion of shift–invariance includes the idea that there is no real time origin, while the TPBVDS (2.1)–(2.2) is defined on an interval [0,N] of fixed length.

We now make the following precise definitions:

Definition 3.1

The TPBVDS (2.1)–(2.3) is left (right) input–output extendible if given any interval [K,N] ([0,L]) containing [0,N], there exists a TPBVDS over this larger interval with the same dynamics as in (2.1) but with new boundary matrices $V_i(K,N)$, $V_f(K,N)$ ($V_i(0,L)$, $V_f(0,L)$) such that the weighting pattern W(k,j) of the original system is the restriction of the weighting pattern $W_e(k,j)$ of the new extended system, i.e.

$$W(k,j) = W_e(k,j), \quad 0 \le k \le N,\ 0 \le j \le N-1. \tag{3.22}$$

The TPBVDS (2.1)–(2.3) is input–output extendible if it is both left and right input–output extendible.

Definition 3.2

The TPBVDS (2.1)–(2.2) is left (right) extendible if given any interval [K,N] ([0,L]) containing [0,N], there exists a TPBVDS over this larger interval with the same dynamics as in (2.1) but with new boundary matrices $V_i(K,N)$, $V_f(K,N)$ ($V_i(0,L)$, $V_f(0,L)$) such that the Green's function G(k,j) of the original system is the restriction of the Green's function $G_e(k,j)$ of the new extended system, i.e.

$$G(k,j) = G_e(k,j), \quad 0\le k\le N,\ 0\le j\le N-1. \tag{3.23}$$

The TPBVDS (2.1)–(2.2) is extendible if it is both left and right extendible.

In order to characterize the conditions under which each of these types of extendibility hold, let us first define two matrices that will appear on several occasions. Specifically, with any matrix F we associate the Drazin Inverse F^D and its invertible modification $\tilde{F}$. To define these, let T be an invertible matrix such that

$$F = T\begin{bmatrix} M & 0 \\ 0 & N \end{bmatrix}T^{-1} \tag{3.24a}$$

where M is invertible and N is nilpotent (e.g. the real Jordan form has this structure). Then

$$F^D = T\begin{bmatrix} M^{-1} & 0 \\ 0 & 0 \end{bmatrix}T^{-1}. \tag{3.24b}$$

$$\tilde{F} = T\begin{bmatrix} M & 0 \\ 0 & N+I \end{bmatrix}T^{-1}. \tag{3.24c}$$

These matrices have a number of important properties:

(i) F^D and $\tilde{F}$ commute with each other and with F.

(ii) If F is invertible, $F^D=F^{-1}$ and $\tilde{F}=F$.

(iii)
$$F^DF = F^D\tilde{F} \tag{3.25}$$

and if μ is the degree of nilpotency of N, i.e. $N^{\mu-1}\neq 0$, $N^{\mu}=0$, then for $k\ge\mu$

$$F^{k+1}F^D = F^k\ ,\ F^k\tilde{F} = F^{k+1}. \tag{3.26}$$

(iv) Let G be any matrix, then the conditon

$$\mathrm{Ker}(F^{\mu}) \subset \mathrm{Ker}(G) \tag{3.27}$$

is equivalent to

$$GF^DF = G. \tag{3.28}$$

(v) If $\mathscr{H}$ is an F–invariant subspace, then $F^DF\mathscr{H} \subset \mathscr{H}$ and is also F–invariant.

(vi) Let $\{E,A\}$ be a regular pencil in standard–form, then

$$EE^D+AA^D-AA^DEE^D = I. \tag{3.29}$$

Properties (i)–(v) can be easily checked. To see why property (vi) is true we need to first pre and post multiply (3.29) by T and T^{-1} chosen such that

$$TET^{-1} = \begin{bmatrix} E_1 & & \\ & E_2 & \\ & & N_e \end{bmatrix}, \; TAT^{-1} = \begin{bmatrix} N_a & & \\ & A_2 & \\ & & A_3 \end{bmatrix} \tag{3.30}$$

where E_1, E_2, A_2 and A_3 are invertible (see Section 6), in which case

$$TE^DT^{-1} = \begin{bmatrix} E_1^{-1} & & \\ & E_2^{-1} & \\ & & 0 \end{bmatrix}, \; TA^DT^{-1} = \begin{bmatrix} 0 & & \\ & A_2^{-1} & \\ & & A_3^{-1} \end{bmatrix}. \tag{3.31}$$

Then clearly

$$TAA^DT^{-1} = \begin{bmatrix} 0 & & \\ & I & \\ & & I \end{bmatrix}, \; TEE^DT^{-1} = \begin{bmatrix} I & & \\ & I & \\ & & 0 \end{bmatrix} \tag{3.32}$$

and

$$TAA^DEE^DT^{-1} = \begin{bmatrix} 0 & & \\ & I & \\ & & 0 \end{bmatrix} \tag{3.33}$$

which imply the desired result.

Note that without loss of generality, it can always be assumed that the E and A matrices of a TPBVDS in normalized form are in the block form (3.30). This can always be achieved by a coordinate transformation. The corresponding boundary matrices, in this coordinate system, must have the following form

$$V_i = \begin{bmatrix} E_1^{-N} & V_{12}^i & V_{13}^i \\ 0 & V_{22}^i & V_{23}^i \\ 0 & V_{32}^i & V_{33}^i \end{bmatrix}, \; V_f = \begin{bmatrix} V_{11}^f & V_{12}^f & 0 \\ V_{21}^f & V_{22}^f & 0 \\ V_{31}^f & V_{32}^f & A_3^{-N} \end{bmatrix}. \tag{3.34}$$

This is because the TPBVDS is supposed to be in normalized form which means that V_i and V_f must satisfy (2.5) and E^N and A^N have the following block structures respectively,

$$\begin{bmatrix} E_1^N & & \\ & E_2^N & \\ & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & \\ & A_2^N & \\ & & A_3^N \end{bmatrix}.$$

Theorem 3.2

A TPBVDS is left extendible if and only if

$$V_i - V_i E^D E = 0 \tag{3.35a}$$

$$V_f - A^D A V_f = 0. \tag{3.35b}$$

It is right extendible if and only if

$$V_i - E^D E V_i = 0 \tag{3.36a}$$

$$V_f - V_f A^D A = 0. \tag{3.36b}$$

It is extendible if and only if

$$V_i - E^D E V_i E^D E = 0 \tag{3.37a}$$

$$V_f - A^D A V_f A^D A = 0. \tag{3.37b}$$

Corollary

For a displacement TPBVDS the following statements are equivalent

(i) The TPBVDS is right extendible.

(ii) It is left extendible.

(iii) It is extendible.

(iv) The following equations hold

$$V_i - V_i E^D E = 0 \tag{3.38a}$$

$$V_f - V_f A^D A = 0. \tag{3.38b}$$

The corollary follows from the theorem because in the displacement case E, E^D, A and A^D commute with V_i and V_f.

Proof of Theorem 3.2

First we show necessity. Let the TPBVDS be left extendible then it must be obtained by moving in the left boundary of another TPBVDS. Then from (3.6) it can be seen that

$$\mathrm{Ker}(V_i) \subset \mathrm{Ker}(E^k) \tag{3.39a}$$

$$\mathrm{Ker}(V_f') \subset \mathrm{Ker}[(A^k)'] \tag{3.39b}$$

where k is the number of steps that the boundary has moved. If k is larger than the maximum of the nilpotency degrees of E and A, then equations (3.39) and (3.35) are of course equivalent. If the system is right extendible then (3.36) can be shown to be true similarly. And of course (3.35) and (3.36) imply (3.37).

To show the sufficiency of (3.35) we need to construct matrices $V_i(K,N)$ and $V_f(K,N)$, for each $K<0$ so that when we move in these boundary matrices to [0,N] we recover V_i and V_f. Assume then that (3.35) holds and let

$$V_i(K,N) = [I-(A^D)^{-K}V_fA^{N-K}](E^D)^{N-K} \tag{3.40a}$$

$$V_f(K,N) = (A^D)^{-K}V_f \tag{3.40b}$$

First we need to make sure that the extended system is in normalized form, i.e.

$$V_i(K,N)E^{N-K}+V_f(K,N)A^{N-K} = I. \tag{3.41}$$

From (3.40) and using the fact that V_i and V_f are in normalized form, we get that

$$V_i(K,N)E^{N-K}+V_f(K,N)A^{N-K} = (I-AA^D)EE^D+AA^D \tag{3.42}$$

which is equal to the identity matrix (see property (vi)).

Now we have to make sure that by moving in $V_i(K,N)$ and $V_f(K,N)$ to $V_i(0,N)$ and $V_f(0,N)$ we recover V_i and V_f. This can be verified by substituting the matrices in (3.40) into (3.19) with K=0, L=J=N and I=K.

The necessity (3.36) for right extendibility can be proven similarly. Specifically we construct right extended matrices as follows

$$V_i(0,L) = (E^D)^{L-N} V_i \tag{3.43a}$$

$$V_f(0,L) = [I-(E^D)^{L-N} V_i E^L](A^D)^L. \tag{3.43b}$$

To see the necessity of (3.37) for extendibility simply note that (3.37) clearly implies (3.35) and (3.36).

Theorem 3.3

Let a TPBVDS be left (right) input–output extendible. Then we can find an equivalent TPBVDS using the freedom in its boundary matrices such that this new TPBVDS is left (right) extendible.

Conversely, every left (right) extendible TPBVDS is left (right) input–output extendible.

Proof

Let a TPBVDS defined over [0,N] be left input–output extendible then there exist a TPBVDS defined over [–n,N] such that when we move in its boundaries to [0,N] we get a TPBVDS with weighting pattern identical to the weighting pattern of our original TPBVDS, possibly with different boundary matrices. This new representation of our TPBVDS is clearly left extendible because it has been obtained by moving in the left boundary of another system n steps. A similar argument can be used for the case of right extendibility.

The converse of the theorem is trivial.

Theorem 3.4

A TPBVDS is left input–output extendible if and only if

$$O_s(V_i - V_iE^DE)R_s = 0 \tag{3.44a}$$

$$O_s(V_f - A^DAV_f)R_s = 0. \tag{3.44b}$$

It is right input–output extendible if and only if

$$O_s(V_i - E^DEV_i)R_s = 0 \tag{3.45a}$$

$$O_s(V_f - V_fA^DA)R_s = 0. \tag{3.45b}$$

It is input–output extendible if and only if

$$O_s(V_i - E^DEV_iE^DE)R_s = 0 \tag{3.46a}$$

$$O_s(V_f - A^DAV_fA^DA)R_s = 0. \tag{3.46b}$$

Corollary

For a stationary TPBVDS the following statements are equivalent

(i) The TPBVDS is right input–output extendible.

(ii) It is left input–output extendible.

(iii) It is input–output extendible.

(iv) The following equations hold

$$O_s(V_i - V_iE^DE)R_s = 0 \tag{3.47a}$$

$$O_s(V_f - V_fA^DA)R_s = 0 \tag{3.47b}$$

Proof of Theorem 3.4

Suppose that the TPBVDS is left input–output extendible, then from Theorem 3.3 there exists a TPBVDS with the same weighting pattern which is left extendible, i.e. there exist matrices V_i^* and V_f^* such that

$$O_sV_i^*R_s = O_sV_iR_s \tag{3.48a}$$

$$O_sV_f^*R_s = O_sV_fR_s \tag{3.48b}$$

and such that they satisfy (3.35).

Notice that (3.48) implies that

$$O_s V_i^* E^D E R_s = O_s V_i E^D E R_s \tag{3.49a}$$

$$O_s A^D A V_f^* R_s = O_s A^D A V_f R_s \tag{3.49b}$$

because of the invariance properties of the strong reachability and observability matrices (see Section 4). Premultiplying and postmultiplying (3.35) (with V replaced by V^*) by O_s and R_s respectively and using (3.48) and (3.49) we obtain (3.44).

Now suppose that (3.44) holds. Let

$$V_i^* = (I-AA^D)(E^D)^N + AA^D V_i E^D E \tag{3.50a}$$

$$V_f^* = AA^D V_f \tag{3.50b}$$

then we have to show that the new system obtained by replacing V_i and V_f with these boundary matrices is just another representation of the original system. First we need to make sure that this new system is in normalized form:

$$V_i^* E^N + V_f^* A^N = [(I-AA^D)(E^D)^N + AA^D V_i E^D E]E^N + AA^D V_f A^N \tag{3.51a}$$

$$= (I-AA^D)(EE^D)^N + AA^D = I. \tag{3.51b}$$

Equation (3.51b) can be checked easily by using the block form (3.30). What remains to be shown is that (3.48) holds for these matrices. Clearly (3.48b) holds from (3.50) and (3.44b). Showing (3.48a) is more complicated and again we suppose that the system is in the block form (3.30), (3.34). The matrix V_i^* in (3.50a), is given by

$$V_i^* = \begin{bmatrix} I & & \\ & 0 & \\ & & 0 \end{bmatrix} \begin{bmatrix} E_1^{-N} & & \\ & E_2^{-N} & \\ & & 0 \end{bmatrix} + \begin{bmatrix} 0 & & \\ & I & \\ & & I \end{bmatrix} \begin{bmatrix} E_1^{-N} & V_{12}^i & V_{13}^i \\ 0 & V_{22}^i & V_{23}^i \\ 0 & V_{32}^i & V_{33}^i \end{bmatrix} \begin{bmatrix} I & & \\ & I & \\ & & 0 \end{bmatrix} =$$

$$\begin{bmatrix} E_1^{-N} & 0 & 0 \\ 0 & V_{22}^i & 0 \\ 0 & V_{32}^i & 0 \end{bmatrix} \tag{3.52}$$

where V_{22}^i and V_{32}^i are (2,2) and (3,2) blocks of V_i.

The strong reachability and observability matrices have a block structure as well , i.e.

$$O_s = W.\begin{bmatrix} O_s^1 & & \\ & O_s^2 & \\ & & O_s^3 \end{bmatrix}, R_s = \begin{bmatrix} R_s^1 & & \\ & R_s^2 & \\ & & R_s^3 \end{bmatrix}.Z \tag{3.53}$$

for some invertible matrices Z and W (this is due to the fact that the three blocks of the system have distinct eigenvalues, see Section 6). Also observe that

$$O_s(V_iE^N+V_fA^N)R_s = O_sV_iE^NR_s + O_sAA^DV_fA^NR_s = O_sR_s. \tag{3.54}$$

By pre and post multiplying (3.54) by W^{-1} and Z^{-1}, respectively, and inspecting the (1,2) block we get that

$$O_s^1V_{12}^iE_2^NR_s^2 = 0 \tag{3.55a}$$

which, since R_s^2 is E_2–invariant (again see Section 6), and E_2 is invertible, implies that

$$O_s^1V_{12}^iR_s^2 = 0. \tag{3.55b}$$

Also note that (3.44a) implies that

$$O_s^kV_{k3}^iR_s^3 = 0, \; k=1,2,3. \tag{3.56}$$

Now by noting the expression for V_i^* in (3.52) and equations (3.55b) and (3.56) it becomes clear that we must have that

$$O_sV_iR_s = O_sV_i^*R_s, \tag{3.57}$$

which is the desired result. The other cases can be argued similarly.

The corollary is immediate by noting that E and A and their Drazin inverses commute with V_i and V_f when premultiplied by O_s and postmultiplied by R_s.

We shall emphasize the fact that the notions of left and right extendibility are indeed distinct notions and we could very well have a system that is right extendible but not left extendible and vice versa. The following example demonstrates this fact:

Example 3.1

Consider the following TPBVDS

$$x(k+1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} x(k) + u(k) \tag{3.58a}$$

$$x(0) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x(N) = v. \tag{3.58b}$$

This TPBVDS is well–posed and in normalized form. It is easy to check that for this system (3.36b) is violated but (3.35a) and (3.35b) hold. Thus this sytem is left extendible but not right extendible.

The input–output extendibility feature is a property of the weighting pattern of the system and not of any specific representation so that it is possible to refer to this property as extendibility of the weighting pattern. The following theorem justifies this:

Theorem 3.5

Let 2 TPBVDS's (of possibly different dimension) defined over [0,N] have identical weighting patterns. Then if one is input–output extendible so is the other.

We shall prove this result in Section 5.

The extendibility property is a very important property because it allows us to associate to each system a sequence of systems defined over any desired interval. We present a way of constructing this sequence. But first we give the following characterization of extendible systems:

Theorem 3.6

Let a TPBVDS be extendible and in block form (3.30), (3.34). Then the boundary matrices must have the following structure

$$V_i = \begin{bmatrix} E_1^{-N} & 0 & 0 \\ 0 & V_{22}^i & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad V_f = \begin{bmatrix} 0 & 0 & 0 \\ 0 & V_{22}^f & 0 \\ 0 & 0 & A_3^{-N} \end{bmatrix} \tag{3.59}$$

which means that the TPBVDS is separated into a purely causal part and a purely anticausal part, each having nilpotent dynamics, and a non–descriptor acausal part.

Note that if a system is input–output extendible then it has a representation of the form (3.59).

Proof

The structure in equation (3.59) can be easily derived from the extendibility condition and the fact that the system is in normalized form.

Theorem 3.6 allows us to simplify the expression for the Green's function solution of an extendible system. By replacing the V_i and V_f in the general Green's function solution by V_i and V_f in (3.59), we obtain the following expression for the Green's function of an extendible system:

$$G(k,j) = \begin{bmatrix} -\tilde{A}^k\tilde{E}^{N-k}V_f EE^D\tilde{E}^{j-N}\tilde{A}^{N-j-1} - (I-EE^D)E^{j-k}(A^D)^{j-k+1} & j\geq k \\ \tilde{A}^k\tilde{E}^{N-k}V_i AA^D\tilde{A}^{-j-1}\tilde{E}^j + (I-AA^D)(E^D)^{k-j}A^{k-j-1} & j<k \end{bmatrix}$$

$$= \begin{bmatrix} -\tilde{A}^k\tilde{E}^{-k}[I-(E^NV_i)]EE^D\tilde{A}^{-j-1}\tilde{E}^j - (I-EE^D)E^{j-k}(A^D)^{j-k+1} & j\geq k \\ \tilde{A}^k\tilde{E}^{-k}(E^NV_i)AA^D\tilde{A}^{-j-1}\tilde{E}^j + (I-AA^D)(E^D)^{k-j}A^{k-j-1} & j<k \end{bmatrix} \tag{3.60}$$

And of course the weighting pattern of an input–output extendible system can be expressed as:

$$W(k,j) = \begin{bmatrix} -C\{\tilde{A}^k\tilde{E}^{-k}[I-(E^NV_i)]EE^D\tilde{A}^{-j-1}\tilde{E}^j - (I-EE^D)E^{j-k}(A^D)^{j-k+1}\}B & j\geq k \\ C\{\tilde{A}^k\tilde{E}^{-k}(E^NV_i)AA^D\tilde{A}^{-j-1}\tilde{E}^j + (I-AA^D)(E^D)^{k-j}A^{k-j-1}\}B & j<k \end{bmatrix} \tag{3.61a}$$

Note that (3.60) expresses the Green's function of an extendible TPBVDS and all of its extension. Similarly, (3.61a) expresses the weighting pattern of an input–output extendible TPBVDS and all of its extensions. This observation deserves further comment. Specifically, what we have done is the following. We begin with a specific extendible TPBVDS defined on [0,N], with boundary matrices V_i, V_f so that the system is in standard form over this specific interval. Equations (3.60) and (3.61a) then provide us with the Green's function and weighting pattern for <u>all</u> extensions of the TPBVDS. Thus we use the parameters associated with any one of the family of extensions to obtain G and W for the whole family. These expressions must of course, not depend on the particular member of the family used in the computation. In particular (3.60) and (3.61a) do not depend on N. Rather $E^N V_i$ is, in a sense, an invariant of the entire family (remember that V_i also depends on N, as it is chosen so that the system is in standard form over [0,N]). In the simpler stationary case this point can be made much more explicit.

Note also that if we are in the basis (3.30), then by letting C and B equal $[C_1 \;\; C_2 \;\; C_3]$, $\begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}$ respectively, W(k,j) can be expressed as

$$W(k,j) = \begin{bmatrix} -C_2 A_2^k E_2^{N-k} V_{22}^f E_2^{j-N} A_2^{N-j-1} B_2 - C_3 A_3^{k-j-1} N_e^{j-k} B_3 & j \geq k \\ C_2 A_2^k E_2^{N-k} V_{22}^i E_2^j A_2^{-j-1} B_2 + C_1 E_1^{j-k} N_a^{k-j-1} B_1 & j < k \end{bmatrix}. \tag{3.61b}$$

We can construct the sequence of (inward and outward) extensions (in standard form) of our extendible or input–ouput extendible TPBVDS as follows

$$\begin{aligned} V_i(I,J) &= \tilde{E}^{-J}\tilde{A}^I (E^N V_i) A A^D \tilde{E}^I \tilde{A}^I + (I - AA^D)\tilde{E}^{I-J}, \\ V_f(I,J) &= \tilde{E}^{-J}\tilde{A}^I [I - (E^N V_i)] E E^D \tilde{E}^J \tilde{A}^{-J} + (I - EE^D)\tilde{A}^{I-J}. \end{aligned} \tag{3.62a}$$

In the basis (3.30), (3.62a) becomes

$$V_i(I,J) = \begin{bmatrix} E_1^{-(J-I)} & 0 & 0 \\ 0 & E_2^{N-J} A_2^I V_{22}^i E_2^I A_2^{-I} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; V_f(I,J) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & A_2^I E_2^{N-J} V_{22}^f E_2^{J-N} A_2^{N-J} & 0 \\ 0 & 0 & A_3^{-(J-I)} \end{bmatrix}. \tag{3.62b}$$

In the stationary case the situation is even simpler. The weighting pattern of an input–output extendible stationary TPBVDS can be expressed as follows

$$W(k) = \begin{cases} C(E^N V_i)E^D(AE^D)^{k-1}B & k>0 \\ -C(A^N V_f)A^D(EA^D)^{-k}B & k\le 0 \end{cases}$$

$$= \begin{cases} C(E^N V_i)E^D(AE^D)^{k-1}B & k>0 \\ C[I-(E^N V_i)]A^D(EA^D)^{-k}B & k\le 0 \end{cases} \quad (3.63)$$

4–Reachability and Observability

As discussed in [1], there are two notions for both reachability and observability for TPBVDS's. In this section we provide brief reviews of these definitions and present additional results.

Definition 4.1

The system (2.1)–(2.2) is strongly reachable on [K,L] if the map

$$\{u(k): k\varepsilon[K,L]\} \rightarrow z_0(K,L) \tag{4.1}$$

is onto. The system is strongly reachable if it is strongly reachable on some interval.

From (3.2) we can write

$$z_0(K,L) = R_s(L-K)\begin{bmatrix} u(K) \\ \vdots \\ u(L-1) \end{bmatrix} \tag{4.2}$$

where

$$R_s(j) = [A^{j-1}B|EA^{j-2}B|\ldots|E^{j-1}B]. \tag{4.3}$$

Note that $R_s=R_s(n)$. Furthermore a TPBVDS is strongly reachable if and only if R_s has full rank (this is a consequence of the generalized Cayley–Hamilton Theorem). Furthermore, the strongly reachable spaces have the usual nesting property, i.e.

$$\mathcal{R}_s(k) = \mathrm{Im}[R_s(k)] \subset \mathrm{Im}[R_s(k+1)] = \mathcal{R}_s(k+1). \tag{4.4}$$

We refer the reader to [1] for proofs of these and other results related to strong reachability. For future reference we define the strongly reachable subspace

$$\mathcal{R}_s = \mathrm{Im}[R_s]. \tag{4.5}$$

Also in [1] it is shown that the system (2.1)–(2.2) is strongly reachable if and only if

$$[sE-tA|B]$$

has full row–rank for all $(s,t)\neq 0$. Or equivalently, if no left eigenvector of the system is orthogonal to the columns of B. Note that this test of strong reachability can be applied even if the system is not in normalized–form. To see this, let us suppose that (2.1)–(2.2) is well–posed, but not in normalized–form. Now remember that any well–posed system can be put into normalized–form by premultiplying (2.1) and (2.2) by some invertible matrices T and S, respectively. Premultiplying (2.1) by an invertible matrix T means replacing E, A, and B with TE, TA and TB, respectively. Since the new system is in normalized form, we can test whether or not it is strongly reachable by testing the full rankedness of $[sTE-tTA|TB]=T[sE-tA|B]$. By noting that T is invertible, we obtain the desired result.

<u>Defintion 4.2</u>

The system (2.1)–(2.3) is <u>strongly observable</u> on [K,L] if the map

$$z_i(K,L) \rightarrow \{y(k): k\varepsilon[K,L]\} \tag{4.6}$$

defined by (3.5), (3.6), and (2.3) with $u\equiv 0$ on [K,L] is one to one. The system is <u>strongly observable</u> if it is strongly observable on some [K,L].

With $u\equiv 0$, we have

$$\begin{bmatrix} y(K) \\ \vdots \\ y(L) \end{bmatrix} = O_s(L-K)z_i(K,L) \tag{4.7}$$

where

$$O_s(j) = \begin{bmatrix} CE^j \\ CAE^{j-1} \\ \vdots \\ CA^j \end{bmatrix}. \tag{4.8}$$

Note that $O_s = O_s(n-1)$. Furthermore, a TPBVDS is strongly observable if and only if O_s has full rank. In addition, the strong unobservability subspaces have the usual nesting property

$$\mathcal{O}_s(k+1) = \mathrm{Ker}(O_s(k+1)) \subset \mathcal{O}_s(k) = \mathrm{Ker}(O_s(k)). \tag{4.9}$$

Again for future reference we define the strongly unobservable subspace

$$\mathcal{O}_s = \mathrm{Ker}(O_s). \tag{4.10}$$

Again form [1] we have that the TPBVDS (2.1)–(2.2) is strongly observable if and only if

$$\begin{bmatrix} sE-tA \\ C \end{bmatrix}$$

has full column–rank for all $(s,t)\neq(0,0)$, or equivalently, if no right eigenvector of the system is orthogonal to the rows of C. Again, as for the case of strong reachability, this test of strong observability is valid even if the system is not in normalized–form.

Note that the properties of strong reachability and observability involve only the matrices C, E, A, and B. As we shall see, the other weaker set of notions of reachability and observability involve the boundary matrices as well.

Definition 4.3

The system (2.1)–(2.2) is weakly reachable off [K,L] if the map

$$\{u(k):\ k\varepsilon[0,K-1]\cup[L,N-1]\} \ \text{->}\ z_i(K,L) \tag{4.11}$$

with $v=0$ is onto. The weakly reachable subspace $\mathcal{R}_w(K,L)$ is the range of this map. The system is called weakly reachable if

$$\mathcal{R}_w \triangleq \bigcup_{K,L} \mathcal{R}_w(K,L) = \mathbb{R}^n. \tag{4.12}$$

The space $\mathcal{R}_w$ is called the weak reachability space.

While it is shown in [1] that for K and L far from the boundaries the dimension of $\mathcal{R}_w(K,L)$ is constant, it is not generally true that this space is fixed or that any nesting of weak reachability spaces occurs as K and L move inward from the boundaries. That is why we may very well have a system which is weakly reachable, but where $\mathcal{R}_w(K,L)$ is not the whole space for any K and L. In [1] we defined weak reachability differently, specifically we called a system weakly reachable if $\mathcal{R}_w(K,L)$

equaled $\mathbb{R}^n$ for K and L far from the boundaries. We shall see later that Definition 4.3 is more appropriate.

Theorem 4.1

The weak reachability space $\mathcal{R}_w$ can be expressed as

$$\mathcal{R}_w = \bigcup_{0\le k<n} A^k E^{n-1-k} \mathrm{Im}[V_i R_s \;\; V_f R_s]. \tag{4.13}$$

Corollary

For an extendible system, the weak reachability space $\mathcal{R}_w$ can be expressed as

$$\mathcal{R}_w = \bigcup_{0\le k<n} A^k E^{n-1-k} V_i \mathcal{R}_s + \mathcal{R}_s. \tag{4.14}$$

Proof

First we prove the following Lemma which justifies the use of the terms "strong" and "weak".

Lemma 4.1

For any TPBVDS

$$\mathcal{R}_s \subset \mathcal{R}_w. \tag{4.15}$$

Proof

We will show a stronger result that

$$\mathcal{R}_s \subset \mathcal{R}_w(K,L) \text{ for } K,L \in [n,N-n]. \tag{4.16}$$

From expression (3.7) for $z_i(K,L)$ with $v=0$, and the fact that the space reached by $z_o(0,K)$ and $z_o(L,N)$ is exactly $\mathcal{R}_s$, we can easily deduce that

$$\mathcal{R}_w(K,L) = E^{N-L}(\omega E - A^K(\omega V_f E + V_i A)A^{N-K})\Gamma^{-1}\mathcal{R}_s$$
$$+ A^K(A - E^{N-L}(\omega V_f E + V_i A)E^L)\Gamma^{-1}\mathcal{R}_s \tag{4.17}$$

By noting that $A^{N-K}\mathcal{R}_s \subset \mathcal{R}_s$ and $E^L\mathcal{R}_s \subset \mathcal{R}_s$, (4.17) implies that

$$\mathcal{R}_w(K,L) \supset E^{N-L}(\omega E - A^K(\omega V_f E + V_i A)A^{N-K})E^L\Gamma^{-1}\mathcal{R}_s$$
$$+ A^K(A - E^{N-L}(\omega V_f E + V_i A)E^L)A^{N-K}\Gamma^{-1}\mathcal{R}_s$$
$$= \mathcal{R}_s + E^{N-L}A^K(\omega V_f E + V_i A)A^{N-K}E^L\Gamma^{-1}\mathcal{R}_s \tag{4.18}$$

which in turn implies (4.16). Clearly then (4.16) implies (4.15).

To prove the Theorem, observe that

$$\mathcal{R}_w \supset \mathcal{R}_s + \bigcup_K \mathcal{R}_w(K,K). \tag{4.19}$$

Using (2.7) and the fact that $\Gamma^{-1}\mathcal{R}_s = \mathcal{R}_s$, this implies that

$$\mathcal{R}_w \supset \bigcup_K \{A^K E^{N-K}[\omega V_f E + V_i A]\mathcal{R}_s\} + \mathcal{R}_s \text{, for all } \omega. \tag{4.20}$$

But thanks to (2.2.5) and the E– and A–invariance of $\mathcal{R}_s$,

$$\mathcal{R}_s + [\omega V_f E + V_i A]\mathcal{R}_s = V_i\mathcal{R}_s + V_f\mathcal{R}_s \tag{4.21}$$

which along with (4.19) and Cayley–Hamilton proves that

$$\mathcal{R}_w \supset \bigcup_{0\le k<n} A^k E^{n-1-k}\mathrm{Im}[V_i R_s \quad V_f R_s]. \tag{4.22}$$

The other inclusion is trivial since in expression (3.7) for $z_i(K,L)$, the range of the map $u \to z_i$ is essentially the range of matrices $A^r E^s V_i A^t E^u$ and $A^r E^s V_f A^t E^u$.

To prove the corollary simply note that we can decompose the system into 3 subsystems as in (3.30), in which case V_i and V_f are expressed as in (3.59). Now using the fact that for an extendible system $V_f = (I - V_i E^N)(A^D)^N$, we can show that

$$\mathrm{Im}[V_i R_s | R_s] = \mathrm{Im}[V_i R_s | V_f R_s] \tag{4.23}$$

which yields the desired result.

In the case of displacement systems, expression (4.13) simplifies and $\mathcal{R}_w$ can be expressed as follows

$$\mathcal{R}_w = \mathrm{Im}[V_i R_s \quad V_f R_s]. \tag{4.24}$$

If in addition the system is extendible then $\mathcal{R}_w = \mathcal{R}_w(K,L)$ for K and L far from the boundaries (see [1]).

In analogy with the strong reachability result, we state without proof the following characterization of weak reachability, which is proved in [2]:

<u>Theorem 4.2</u>

A displacement system is weakly reachable if and only if the matrix

$$[sE-tA|V_iB|V_fB]$$

has full rank for all $(s,t)\neq(0,0)$. If in addition it is extendible, then it is weakly reachable if and only if

$$[sE-tA|V_iB|B]$$

has full rank for all $(s,t)\neq(0,0)$.

As one would expect, there is a dual set of concepts and results for weak observability:

<u>Definition 4.4</u>

The system (2.1)–(2.3) is <u>weakly observable</u> off [K,L] if the map

$$z_o(K,L) \rightarrow \{y(k): k\varepsilon[0,K]\cup[L,N]\} \tag{4.25}$$

with v=0 and $u(j)=0$, $j\varepsilon[0,K-1]\cup[L,N-1]$ is one to one. The weakly unobservable subspace $\mathcal{O}_w(K,L)$ is the null space of this map. The system is called weakly observable if

$$\mathcal{O}_w \triangleq \bigcap_{K,L} \mathcal{O}_w(K,L) = \{0\}. \tag{4.26}$$

The space $\mathcal{O}_w$ is the weak unobservability space.

By analogy with the weak reachability case we simply present the dual set of results concerning weak observability.

Theorem 4.3

The weakly unobservable space can be expressed as follows

$$\mathcal{O}_w = \bigcap_{0\le k<n} \operatorname{Ker} \begin{bmatrix} O_s V_i \\ O_s V_f \end{bmatrix} E^{n-1-k} A^k. \tag{4.27}$$

Corollary

For an extendible system the weakly unobservable space can be expressed as follows

$$\mathcal{O}_w = \mathcal{O}_s \cap \{ \bigcap_{0\le k<n} \operatorname{Ker}(O_s V_i E^{n-1-k} A^k) \}. \tag{4.28}$$

Lemma 4.2

For any TPBVDS

$$\mathcal{O}_w \subset \mathcal{O}_s. \tag{4.29}$$

This Lemma shows that weak observability is a weaker condition that strong observability.

If the TPBVDS is displacement, (4.27) simplifies as follows

$$\mathcal{O}_w = \operatorname{Ker} \begin{bmatrix} O_s V_i \\ O_s V_f \end{bmatrix} \tag{4.30}$$

and if in addition it is extendible $\mathcal{O}_w = \mathcal{O}_w(K,L)$ for all K and L far from the boundaries.

Theorem 4.4

A displacement system is weakly observable if and only if the matrix

$$\begin{bmatrix} sE-tA \\ CV_i \\ CV_f \end{bmatrix}$$

has full rank for all $(s,t)\neq(0,0)$. If in addition, it is extendible, then it is extendible if and only if

$$\begin{bmatrix} sE-tA \\ CV_i \\ C \end{bmatrix}$$

has full rank for all $(s,t)\neq(0,0)$.

In this section we have introduced two distinct notions of reachability and observability. In some cases the two notions coincide, for example for causal systems where $V_i=E=I$ and $V_f=0$. In this case $\mathcal{R}_s=\mathcal{R}_w$ and $\mathcal{O}_s=\mathcal{O}_w$. In general, however, that is not the case. We shall see in the next section that both of these notions are indeed needed to study minimality.

The following example illustrates the difference between the concept of strong and that of weak reachability:

Example 4.1

Consider the following displacement TPBVDS

$$x(k+1) = x(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(k) \tag{4.31}$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} x(0) + \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} x(N) = 0 \tag{4.32}$$

where

$$x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}. \tag{4.33}$$

This system is well–posed and in normalized form. The strong reachability space for this system is just

$$\mathrm{Im}\begin{bmatrix}1\\0\end{bmatrix},$$

so that the system is not strongly reachable. In fact, we can easily see that only x_1 is strongly reachable and x_2 is not. However, using (4.24), we can check that the system is weakly reachable. In fact, we can check that this system is weakly reachable off any interval [K,L], 0<K,L<N. To intuitively illustrate this fact, note that boundary condition (4.32) can be rewritten as follows

$$x_1(0) = 0 \tag{4.34}$$

$$x_2(0) = x_1(N). \tag{4.35}$$

It is clear that $x_1(k)$ can be made arbitrary by proper choice of inputs $u(j)$, $j<k$. On the other hand, $x_1(N)$, and thus $x_2(0)$, can also be independently made arbitrary by proper choice of $u(j)$, $k\leq j<N$. But (4.31) implies that $x_2(k)$ is constant for all k, so that it must equal $x_2(0)$ and $x_1(N)$. The result is that $x_1(k)$ and $x_2(k)$, which form $x(k)$, can be made arbitrary by proper choice of the input u. This explains why this system is weakly reachable.

5–Minimality

In this section we present minimality results for TPBVDS's. We also specifically consider the stationary and extendible stationary cases. These results are analogous to those in [7] and [12], with differences due to possible singularity of E and A.

Definition 5.1

A TPBVDS is minimal if x has the lowest dimension among all TPBVDS's having the same weighting pattern.

Theorem 5.1

A TPBVDS with $N \geq 4n$ is minimal if and only if

(a) $\mathcal{R}_w = \mathbb{R}^n$ (5.1)

(b) $\mathcal{O}_w = \{0\}$ (5.2)

(c) $\mathcal{O}_s \subset \mathcal{R}_s$ (5.3)

(i.e. if it is weakly reachable and observable, and any strongly unobserved mode is strongly reached).

As in the causal case, the proof of this result involves the introduction of Hankel matrices and the description of a method for reducing the dimension of systems violating any of the conditions (a)–(c). As we will see, in the present context we actually have 3 different Hankel matrices and also, as in [7] we may have a certain level of nonuniqueness in minimal realizations that is not present in the causal case.

The length of the interval here is assumed to be larger than 4 times the dimension of the system so that all the modes on both sides of a state in the middle of the interval can be reached and observed (see the proof for details on where this assumption is needed). If N is not large enough, the conditions of Theorem 5.1 become necessary but not sufficient.

Proof

We begin with the description of reduction procedures if any of the conditions (5.1)–(5.3) are not satisfied. Consider first the case in which $\mathcal{R}_w \neq \mathbb{R}^n$. Let $\mathcal{R}_2$ be any subspace such that

$$\mathcal{R}_w \oplus \mathcal{R}_2 = \mathbb{R}^n. \tag{5.4}$$

Then, by performing a similarity transformation on x to represent it in a basis compatible with (5.4) we arrive at a system of the form (1)–(3) with

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad x_1 \in \mathcal{R}_w,\ x_2 \in \mathcal{R}_2, \tag{5.5a}$$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},\ E = \begin{bmatrix} E_{11} & E_{12} \\ 0 & E_{22} \end{bmatrix},\ V_i = \begin{bmatrix} V^i_{11} & V^i_{12} \\ V^i_{21} & V^i_{22} \end{bmatrix}, \tag{5.5b}$$

$$V_f = \begin{bmatrix} V^f_{11} & V^f_{12} \\ V^f_{21} & V^f_{22} \end{bmatrix},\ C = [C_1 \mid C_2],\ B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}. \tag{5.5c}$$

The 0–blocks in A and E follow from the A– and E–invariance of $\mathcal{R}_w$. The 0–block in B is due to the fact that $\mathrm{Im}[B] \subset \mathcal{R}_s$ (Cayley–Hamilton) and $\mathcal{R}_s \subset \mathcal{R}_w$. In addition, since

$$\mathcal{R}_w \supset V_i \mathcal{R}_s + V_f \mathcal{R}_s \tag{5.6}$$

we must have

$$V^i_{21} A^k_{11} E^j_{11} B_1 = V^f_{21} A^k_{11} E^j_{11} B_1 = 0. \tag{5.7}$$

From the form (2.7) for the weighting pattern of a TPBVDS we can then conclude that the weighting pattern of our system is given by

$$W(k,j) = \begin{cases} C_1 A^k_{11}(A_{11} - E^{N-k}_{11}(V^i_{11}A_{11} + \omega V^f_{11}E_{11})E^k_{11})E^{j-k}_{11} A^{N-j-1}_{11} \Gamma^{-1}_1 B_1 & j \geq k \\ C_1 E^{N-k}_{11}(\omega E_{11} - A^k_{11}(V^i_{11}A_{11} + \omega V^f_{11}E_{11})A^{N-k}_{11})E^j_{11} A^{k-j-1}_{11} \Gamma^{-1}_1 B_1 & j < k \end{cases} \tag{5.8}$$

where $\Gamma_1 = \omega E_{11}^{N+1} - A_{11}^{N+1}$, so that we have apparently reduced our system to

$$\begin{aligned} E_{11}x_1(k+1) &= A_{11}x_1(k) + B_1u(k) \\ V_{11}^i x_1(0) + V_{11}^f x_1(N) &= v_1 \\ y(k) &= C_1x_1(k). \end{aligned} \tag{5.9}$$

Note since E and A are in standard form, so are E_{11} and A_{11}. However, the boundary matrices V_{11}^i and V_{11}^f need not be normalized and indeed there is no guarantee that (5.9) is well–posed. However, note that V_{11}^i and V_{11}^f in (5.8) are pre– and post–multiplied by $C_1A_{11}^kE_{11}^{N-k}$ and $\Gamma_1^{-1}A_{11}^jE_{11}^{N-j}B_1$ so that there may be a degree of freedom in the choices for V_{11}^i and V_{11}^f that verify (5.8). In fact, if O_s^1 and R_s^1 denote the strong observability and reachability matrices of (5.9), since the null space of $C_1A_{11}^kE_{11}^{N-k}$ includes that of O_s^1 and the range of $\Gamma_1^{-1}A_{11}^jE_{11}^{N-j}B_1$ is included in R_s^1, we can modify V_{11}^i and V_{11}^f as long as $O_s^1V_{11}^iR_s^1$ and $O_s^1V_{11}^fR_s^1$ remain unchanged and preserve the weighting pattern (5.8). Thanks to the following result, we can modify the boundary matrices in order to make (5.9) well–posed while leaving the weighting pattern unchanged.

Lemma 5.1

Consider a (possibly not well–posed) TPBVDS (2.1)–(2.2), with E and A in standard form and for which the following holds:

$$O_s(V_iE^N + V_fA^N)R_s = O_sR_s \tag{5.10}$$

Then we can find $\tilde{V}_i$, $\tilde{V}_f$ so that

$$\tilde{V}_iE^N + \tilde{V}_fA^N = I \tag{5.11}$$

and

$$O_sV_iR_s = O_s\tilde{V}_iR_s \ , \ O_sV_fR_s = O_s\tilde{V}_fR_s. \tag{5.12}$$

Proof

Let

$$X = I - [V_iE^N + V_fA^N] \tag{5.13}$$

so that

$$O_s X R_s = 0. \tag{5.14}$$

Let a and b be any scalars such that (aE^N+bA^N) is invertible, and then take

$$\tilde{V}_i = V_i + aX(aE^N + bA^N)^{-1} \tag{5.15a}$$

$$\tilde{V}_f = V_f + bX(aE^N + bA^N)^{-1} \tag{5.15b}$$

From (5.14) and the A– and E–invariance of $\mathcal{R}_s$ and $\mathcal{O}_s$ we have that

$$O_s A^k E^j X A^r E^s R_s = 0 \qquad k,j,r,s \geq 0 \tag{5.16}$$

from which we can easily check that (5.12) holds. Finally (5.11) can also be checked by direct calculation.

To apply this lemma to (5.9) we must show that (5.10) holds for this system. Expressions (5.5) and (5.7) imply that:

$$CA^k E^j B = C_1 A_{11}^k E_{11}^j B_1 \tag{5.17a}$$

$$CA^r E^s V_i A^k E^j B = C_1 A_{11}^r E_{11}^s V_{11}^i A_{11}^k E_{11}^j B_1 \tag{5.17b}$$

$$CA^r E^s V_f A^k E^j B = C_1 A_{11}^r E_{11}^s V_{11}^f A_{11}^k E_{11}^j B_1. \tag{5.17c}$$

Therefore, since our original system was assumed to be in normalized form

$$\begin{aligned} C_1 A^r E^s [V_{11}^i E_{11}^N + V_{11}^f A_{11}^N] A_{11}^k E_{11}^j B_1 &= CA^r E^s [V_i E^N + V_f A^N] A^k E^j B \\ &= CA^{k+r} E^{j+s} B = C_1 A_{11}^{k+r} E_{11}^{j+s} B_1 \end{aligned} \tag{5.18}$$

from which we conclude that

$$O_s^1 [V_{11}^i E_{11}^N + V_{11}^f A_{11}^N] R_s^1 = O_s^1 R_s^1 \tag{5.19}$$

where O_s^1 and R_s^1 are the strong observability and reachability matrices for (5.9).

To continue with the proof of the theorem, note that the problem of reducing the dimension of the realization if (5.2) is violated is merely the dual of the problem that we have just considered. Consequently, we omit the details. We turn then to the case in which condition (5.3) is not satisfied. In this case, there is a subspace $\mathcal{Z} \neq \{0\}$ such that

$$\mathcal{R}_s \oplus \mathcal{Z} = \mathcal{R}_s + \mathcal{O}_s. \tag{5.20}$$

Let $\mathscr{W}$ be any subspace such that $\mathscr{W} \oplus \mathscr{Z} = \mathbb{R}^n$ and perform a similarity transformation of the TPBVDS to represent it in a basis compatible with (5.4). This yields a model as in (5.5b), (5.5c) with the additional fact that $C_2=0$. To put the reduced system in normalized form we once again apply Lemma 5.1.

What remains to show is that two TPBVDS's with the same weighting pattern and both satisfying (5.1)–(5.3) must have the same dimension and consequently are minimal. To proceed with the proof we need the following Lemma:

Lemma 5.2

Let $\{E_i,A_i\}$, i=1,2, be two regular pencils so that $\alpha E_i+\beta A_i=I$, i=1,2, where $\dim(E_i)=\dim(A_i)=n_i$. Suppose that $N\geq 2\max(n_1,n_2)$. Also suppose that for some matrices $\{M_i,N_i\}$, i=1,2,

$$M_1A_1^kE_1^{N-1-k}N_1 = M_2A_2^kE_2^{N-1-k}N_2\ , \quad 0\leq k\leq N-1. \tag{5.21}$$

Then for all K, L,

$$M_1A_1^KE_1^LN_1 = M_2A_2^KE_2^LN_2. \tag{5.22}$$

Proof

Note first that for $K+L\leq N-1$ we can write

$$E_i^KA_i^L = E_i^KA_i^L(\alpha E_i+\beta A_i)^{N-1-K-L} \tag{5.23}$$

and in this case (5.22) follows directly from (5.21). For $K+L\geq N$, let us suppose for simplicity that $\alpha\neq 0$. From what we have first shown for $K+L\leq N-1$, we know that

$$M_1A_1^kN_1 = M_2A_2^kN_2\ , \quad 0\leq k\leq N-1. \tag{5.24}$$

From results on the causal partial realization problem [15] and the fact that $N\geq 2n_i$, we can conclude that

$$M_1A_1^kN_1 = M_2A_2^kN_2\ , \quad k\geq 0. \tag{5.25}$$

Equation (5.22) then follows since we can write E_i as $(I-\beta A_i)/\alpha$.

We note that as discussed in [15] the condition on the size of N is important here, although slightly smaller bounds on the interval size can be obtained —(essentially the sum of observability and reachability indices).

Proceeding with the proof, consider two systems $(C_j,E_j,A_j,V_j^i,V_j^f,B_j)$, j=1,2, satisfying minimality conditions (5.1)–(5.3), and without loss of generality assume that both are in normalized form with the same α and β. What we know is that

$$C_1A_1^k(A_1-E_1^{N-k}(V_1^iA_1+\omega V_1^fE_1)E_1^k)E_1^{j-k}A_1^{N-j-1}\Gamma_1^{-1}B_1$$
$$= C_2A_2^k(A_2-E_2^{N-k}(V_2^iA_2+\omega V_2^fE_2)E_2^k)E_2^{j-k}A_2^{N-j-1}\Gamma_2^{-1}B_2\ ,\ j\geq k \tag{5.26a}$$
$$C_1E_1^{N-k}(\omega E_1-A_1^k(V_1^iA_1+\omega V_1^fE_1)A_1^{N-k})E_1^jA_1^{k-j-1}\Gamma_1^{-1}B_1$$
$$= C_2E_2^{N-k}(\omega E_2-A_2^k(V_2^iA_2+\omega V_2^fE_2)A_2^{N-k})E_2^jA_2^{k-j-1}\Gamma_2^{-1}B_2\ ,\ j<k. \tag{5.26b}$$

Let $k\in[2n,N-2n]$ (remember that $N\geq 4n$), then we can apply Lemma 3.2 to get

$$C_1A_1^k(A_1-E_1^{N-k}(V_1^iA_1+\omega V_1^fE_1)E_1^k)E_1^KA_1^L\Gamma_1^{-1}B_1$$
$$= C_2A_2^k(A_2-E_2^{N-k}(V_2^iA_2+\omega V_2^fE_2)E_2^k)E_2^KA_2^L\Gamma_2^{-1}B_2\ ,\ \text{for all } K,L \tag{5.27a}$$
$$C_1E_1^{N-k}(\omega E_1-A_1^k(V_1^iA_1+\omega V_1^fE_1)A_1^{N-k})E_1^KA_1^L\Gamma_1^{-1}B_1$$
$$= C_2E_2^{N-k}(\omega E_2-A_2^k(V_2^iA_2+\omega V_2^fE_2)A_2^{N-k})E_2^KA_2^L\Gamma_2^{-1}B_2\ ,\ \text{for all } K,L. \tag{5.27b}$$

By taking K=r, L=N–k+s in (5.27a) and K=k+r and L=s in (5.27b) and subtracting the two sides of (5.27a) from (5.27b) we obtain

$$C_1E_1^rA_1^sB_1 = C_2E_2^rA_2^sB_2 \tag{5.28}$$

and this for all $r,s\geq 0$.

Using (5.26) and (5.28) we can show that

$$C_1A_1^kE_1^{N-k}(V_1^iA_1+\omega V_1^fE_1)E_1^jA_1^{N-j-1}\Gamma_1^{-1}B_1 = C_2A_2^kE_2^{N-k}(V_2^iA_2+\omega V_2^fE_2)E_2^jA_2^{N-j-1}\Gamma_2^{-1}B_2, \tag{5.29}$$

and taking into account Lemma 5.2, this implies that

$$C_1A_1^rE_1^s(V_1^iA_1+\omega V_1^fE_1)E_1^tA_1^u\Gamma_1^{-1}B_1 = C_2A_2^rE_2^s(V_2^iA_2+\omega V_2^fE_2)E_2^tA_2^u\Gamma_2^{-1}B_2 \tag{5.30}$$

for all $r,s,t,u\geq 0$. Then using the fact that both systems are in normalized form we obtain

$$C_1A_1^rE_1^sV_1^iE_1^tA_1^uB_1 = C_2A_2^rE_2^sV_2^iE_2^tA_2^uB_2 \tag{5.31a}$$

$$C_1A_1^rE_1^sV_1^fE_1^tA_1^uB_1 = C_2A_2^rE_2^sV_2^fE_2^tA_2^uB_2 \tag{5.31b}$$

for all r,s,t,u≥0.

As in the case of causal systems, Hankel matrices are extremely useful in proving our minimality result. In the present context, however, there are three different Hankel matrices.

$$H_{in} = O_s^1R_w^1 = O_s^2R_w^2 \tag{5.32}$$

$$H_{out} = O_w^1R_s^1 = O_w^2R_s^2 \tag{5.33}$$

$$H_s = O_s^1R_s^1 = O_s^2R_s^2 \tag{5.34}$$

where R_s^j and O_s^j are the strong reachability and observability matrices of system j, respectively, and where

$$R_w^j = [A^{n-1}(V_iR_s^j|V_fR_s^j)\ \ EA^{n-2}(V_iR_s^j|V_fR_s^j)\ \ \ldots\ \ E^{n-1}(V_iR_s^j|V_fR_s^j)],$$

$$O_w^j = \begin{bmatrix} \begin{bmatrix} O_s^jV_i \\ O_s^jV_f \end{bmatrix} A^{n-1} \\ \vdots \\ \begin{bmatrix} O_s^jV_i \\ O_s^jV_f \end{bmatrix} E^{n-1} \end{bmatrix},\quad j=1,$$

R_w^j and O_w^j are the <u>weak reachability</u> and <u>weak observability matrices</u> of system j, respectively. Clearly

$$\mathcal{R}_w^j = \mathrm{Im}(R_w^j) \tag{5.35a}$$

$$\mathcal{O}_w^j = \mathrm{Ker}(O_w^j) \tag{5.35b}$$

for j=1, Equations (5.32)–(5.34) are direct consequences of (5.28) and (5.31). From (5.33), we get that

$$R_s^2 = UR_s^1, \tag{5.36}$$

where

$$U = \{O_w^{2\prime}O_w^2\}^{-1}O_w^{2\prime}O_w^1, \tag{5.37}$$

and where O_w^2 has full rank because of the weak observability assumption. Similarly we can obtain an analogous expression for R_s^1 in terms of R_s^2. These allow us to conclude that

$$\mathrm{rank}(R_s^1) = \mathrm{rank}(R_s^2) = \rho, \tag{5.38}$$

and in an analogous way we can show that

$$\mathrm{rank}(O_s^1) = \mathrm{rank}(O_s^2) = \omega. \tag{5.39}$$

Finally, condition (5.3) together with (5.34) imply that

$$\rho - (n_1 - \omega) = \text{rank } H_s = \rho - (n_2 - \omega) \tag{5.40}$$

from which we see that

$$n_1 = n_2, \tag{5.41}$$

completing the proof of the Theorem.

Corollary 5.1a

Let $(C_j, V_j^i, V_j^f, E_j, A_j, B_j, N)$, j=1,2, be two minimal realizations of the same weighting pattern, where $\{E_j, A_j\}$, j=1,2, are in standard form for the same α and β. Then there exists an invertible matrix T so that

$$B_2 = TB_1 \tag{5.42a}$$

$$C_2 = C_1T^{-1} \tag{5.42b}$$

$$O_s^1(V_1^i - T^{-1}V_2^iT)R_s^1 = 0 \tag{5.43a}$$

$$O_s^1(V_1^f - T^{-1}V_2^fT)R_s^1 = 0 \tag{5.43b}$$

and

$$(A_1 - T^{-1}A_2T)R_s^1 = 0 \tag{5.44a}$$

$$(E_1 - T^{-1}E_2T)R_s^1 = 0 \tag{5.44b}$$

$$O_s^1(A_1 - T^{-1}A_2T) = 0 \tag{5.44c}$$

$$O_s^1(E_1 - T^{-1}E_2T) = 0 \tag{5.44d}$$

where R_s^1 and O_s^1 are the strong reachability and observability matrices for system 1.

Proof

From (5.36) we have that

$$R_s^2 = UR_s^1 \tag{5.45}$$

with U defined as in (5.37). While this choice for U is not necessarily invertible, we can always find an invertible T so that

$$R_s^2 = TR_s^1 \tag{5.46}$$

since $\mathcal{R}_s^1$ and $\mathcal{R}_s^2$ have the same dimension. In a similar way we can always find an invertible matrix W so that

$$O_s^2 W = O_s^1. \tag{5.47}$$

From (5.34) we can then conclude that

$$O_s^2 [W-T] R_s^1 = 0. \tag{5.48}$$

The question, then is whether we can choose W=T. To see that this can be done, assume that we have chosen a basis for each of the two systems compatible with the following direct sum decomposition:

$$\mathcal{O}_s \oplus [\mathcal{O}_s^{\perp} \cap \mathcal{R}_s] \oplus [\mathcal{O}_s^{\perp} \cap \mathcal{R}_s].$$

The requirement (5.46) implies that T must have the form

$$T = \begin{bmatrix} T_1 & T_2 & * \\ T_3 & T_4 & * \\ 0 & 0 & * \end{bmatrix} \tag{5.49}$$

where T_1, T_2, T_3 and T_4 are fixed and * are arbitrary. Similarly, (5.47) implies that W must have the form

$$W = \begin{bmatrix} * & * & * \\ 0 & W_1 & W_2 \\ 0 & W_3 & W_4 \end{bmatrix}. \tag{5.50}$$

Finally, by direct computation we can check that (5.48) implies

$$W_1 = T_4, \quad T_3 = W_3 = 0 \tag{5.51}$$

so that with the indicated degrees of freedom we can take

$$W = T = \begin{bmatrix} T_1 & T_2 & * \\ 0 & T_4 & W_2 \\ 0 & 0 & W_4 \end{bmatrix}. \tag{5.52}$$

Proceeding with the proof, note that (5.42a), (5.42b) follow from (5.47), (5.48) plus the fact that $\{E_j, A_j\}$, j=1,2, are in standard form for the same α and β. Also, the equality of the weighting patterns of the two systems is equivalent to

$$O_s^1 V_1^i R_s^1 = O_s^2 V_2^i R_s^2 \qquad (5.53a)$$
$$O_s^1 V_1^f R_s^1 = O_s^2 V_2^f R_s^2 \qquad (5.53b)$$

from which (5.43a) and (5.43b) follow. Finally, recall that R_s is A– and E–invariant. Thus, thanks to Cayley–Hamilton we can conclude that

$$A_2 R_s^2 = T A_1 R_s^1 \ , \quad E_2 R_s^2 = T E_2 R_s^2 \qquad (5.54)$$

from which (5.44a) and (5.44b) follow. Equation (5.44c) and (5.44d) are verified in a similar fashion.

Corollary 5.1b

(a) Every left (right) input–output extendible TPBVDS has a minimal realization that is also left (right) input–output extendible.

(b) Every left (right) extension of a minimal left (right) input–output extendible TPBVDS is minimal.

Proof

Part (a) follows Theorem 3.5 which we prove here.

Proof of Theorem 3.5

Suppose that we have two realizations $(C_j, E_j, A_j, V_j^i, V_j^f, B_j)$, j=1,2, of the same weighting pattern. Then we would like to show that if one of these is left (right) input–output extendible, so is the other. This result can be seen to be true as follows. First, it is not difficult to see that the following genralization of Lemma 5.2 holds. Specifically, if (5.21) holds, then for all P,Q,K,L≥0,

$$M_1 (A_1^D)^P (E_1^D)^Q A_1^K E_1^L N_1 = M_2 (A_2^D)^P (E_2^D)^Q A_2^K E_2^L N_2. \qquad (5.55)$$

Then not only do we have that

$$O_s^1 V_1^i R_s^1 = O_s^2 V_2^i R_s^2 \qquad (5.56a)$$
$$O_s^1 V_1^f R_s^1 = O_s^2 V_2^f R_s^2 \qquad (5.56b)$$

(since both systems have the same weighting pattern) but also

$$O_s^1 V_1^i E_1^D E_1 R_s^1 = O_s^2 V_2^i E_2^D E_2 R_s^2 \qquad (5.57a)$$
$$O_s^1 A_1^D A V_1^f R_s^1 = O_s^2 A_2^D A_2 V_2^f R_s^2. \qquad (5.57b)$$

Suppose that system 1 is left input–output extendible and thus satisfies (3.44). Then (5.56) and (5.57) imply that system 2 also satisfies (3.44) which means that it is left input–output extendible. Right extendibility can be proven similarly.

To show part (b), suppose that an extension of a minimal system defined on the interval [0,N] is not minimal and thus can be reduced. Reduce the extension and move in its boundaries to the interval [0,N]. The system obtained has clearly the same weighting pattern as the original system defined on [0,N] but has lower dimension, which is a contradiction.

Theorem 5.2

A stationary TPBVDS, with $N \geq 2n$, is minimal if and only if

$$\text{(a) } \mathrm{Im}[V_i R_s | V_f R_s] = \mathbb{R}^n \tag{5.58}$$

$$\text{(b) } \mathrm{Ker}\begin{bmatrix} O_s V_i \\ O_s V_f \end{bmatrix} = \{0\} \tag{5.59}$$

$$\text{(c) } \mathcal{O}_s \subset \mathcal{R}_s \tag{5.60}$$

Proof

First, note that the minimality conditions of Theorem 5.1 are necessary and sufficient for this case as well, even though we have a weaker condition on the length of the interval. This is because the only place that the assumption $N \geq 4n$ was used in the proof of Theorem 5.1, was in the derivation of (5.28) and (5.31). But in the stationary case, as long as $N \geq 2n$, (5.31) immediately follows from Lemma 5.2 and the assumption that the weighting patterns of the two systems must be identical. In addtion, (5.28) follows from (5.31) and the assumption that the two systems are in normalized form. So all we need to show is that conditions (5.1)–(5.3) and (5.58)–(5.60) are equivalent in the stationary case.

Note that since

$$\mathrm{Im}[V_iR_s|V_fR_s] \subset \mathcal{R}_w \tag{5.61}$$

$$\mathrm{Ker}\begin{bmatrix} O_sV_i \\ O_sV_f \end{bmatrix} \supset \mathcal{O}_s \tag{5.62}$$

condition (5.58)–(5.60) are sufficient for minimality. To show necessity let us assume that (5.1)–(5.3) hold. Suppose also that (5.58) fails in which case there exists a vector $q \neq 0$ such that

$$q'[V_iR_s|V_fR_s] = 0, \tag{5.63}$$

which implies that

$$q'R_s = 0. \tag{5.64}$$

By noting that condition (5.60) is equivalent to

$$\text{Left–Ker}(R_s) = \mathcal{R}_s^{\perp} \subset \mathcal{O}_s^{\perp} = \text{Row–Im}(O_s) \tag{5.65}$$

(5.64) implies that

$$q' \in \text{Row–Im}(O_s) \tag{5.66}$$

which thanks to the stationarity conditions (2.13a) and (2.13b) implies that

$$q'[V_iE^rA^s - E^rA^sV_i]R_s = 0 \tag{5.67a}$$

$$q'[V_fE^rA^s - E^rA^sV_f]R_s = 0 \tag{5.67b}$$

for all r and s. Thanks to E and A–invariance of R_s, there exists a matrix D such that

$$E^{n-1-k}A^kR_s = R_sD. \tag{5.68}$$

Then (5.67) implies

$$q'E^{n-1-k}A^k[V_iR_s|V_fR_s] = q'[V_iR_sD|V_fR_sD] = 0. \tag{5.69}$$

Since (5.69) holds for all $k \in [0,n-1]$ we obtain

$$q'R_w = 0, \tag{5.70}$$

which violates (5.1). Similarly we can show that if (5.59) fails, then (5.2) is violated.

We have shown above that conditions (5.1)–(5.3) are equivalent to conditions (5.58)–(5.60) for stationary systems. However, note that this does not imply that (5.1) is equivalent to (5.58), and (5.2) to (5.59). As can be seen from the proof of Theorem 5.2, condition (5.60) must be true to have (5.1) be equivalent to (5.58) and for (5.2) to be equivalent to (5.59). The following example illustrates this point:

Example 5.1

Consider the following stationary TPBVDS in normalized form defined over an interval of length N

$$C = [0\ 0\ 1],\ V_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\ V_f = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & N & -N^2/2 \end{bmatrix},$$

$$E = I,\ A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{5.71}$$

For this system, the strong reachability space R_s is

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

and so $[V_iR_s|V_fR_s]$ is equal to

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & -1 \end{bmatrix}$$

and thus does not have full rank. This implies that condition (5.58) is not satisfied. On the other hand, we have that

$$\begin{aligned} \mathcal{R}_w &= \bigcup_{k=0,..,n-1} \{\mathrm{Im}(E^{n-k-1}A^k[V_iR_s|V_fR_s])\} \\ &= \bigcup_{k=0,1,2} \{\mathrm{Im}(\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^k \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & -1 \end{bmatrix})\} = \mathbb{R}^3 \end{aligned} \tag{5.72}$$

which means that condition (5.1) is satisfied. This example illustrates that, if (5.3) does not hold, (5.1) and (5.58) are not equivalent. In this example, (5.3) does not hold since the strong observability matrix O_s is equal to [0 0 1] which implies that $\mathcal{O}_s$ equals

$$\mathrm{Im}(\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix})$$

6–Block Standard and Normalized Forms

A well–known result for causal systems is the following. Suppose that the A–matrix is block diagonalized with no common eigenvalues among the blocks. Then reachability and observability of the entire system is equivalent to the reachability and observability of all of the individual subsystems defined by the block structure of A. The same type of result is easily shown to hold as well for stationary TPBVDS's once we define generalized notions of standard and normalized form.

Definition 6.1

The regular pencil $\{E,A\}$ is in block standard form if

(i) for some invertible matrix T, we have

$$TET^{-1} = \mathrm{diag}(E_1,E_2,\ldots,E_M) \quad (6.1)$$

$$TAT^{-1} = \mathrm{diag}(A_1,A_2,\ldots,A_M) \quad (6.2)$$

where

(ii) each $\{E_i,A_i\}$ pair is in standard form, i.e. there exist α_i, β_i such that

$$\alpha_i E_i + \beta_i A_i = I \quad , \quad i=1,\ldots,M \quad (6.3)$$

and furthermore $\{E_i,A_i\}$ and $\{E_j,A_j\}$, $i\neq j$, have no eigenmode in common. That is, for any pair $(s,t)\neq(0,0)$, $|sE_i-tA_i| = 0$ for at most one value of i.

If a system is in block standard form, we can always, by a change of the coordinate system, transform it such that E and A are block diagonal as in (6.1) and (6.2). Thus for simplicity we can assume that if the system is in block standard form, E and A are block diagonal.

Note that E and A in (6.1), (6.2) commute, and from the proof of the well–posedness result in [1], we can readily check that well–posedness of (2.1)–(2.3) when E and A commute is equivalent to the invertibility of $V_iE^N+V_fA^N$. Consequently, if this is true we can premultiply (2.2) by the inverse of this matrix to obtain a generalization of normalized form:

Definition 6.2

The TPBVDS (2.1)–(2.2) is in block normalized form (BNF) if {E,A} is in block standard form and (2.5) holds.

In general, there is no reason for V_i and V_f to be block–diagonal for a system in BNF. However, in the stationary case we have the following result:

Theorem 6.1

A TPBVDS in BNF is stationary if and only if it has a representation where V_i and V_f are in the same block diagonal form as E and A, i.e.

$$TV_iT^{-1} = \mathrm{diag}(V_1^i,...,V_M^i) \tag{6.4}$$

$$TV_fT^{-1} = \mathrm{diag}(V_1^f,...,V_M^f) \tag{6.5}$$

and moreover, each of the subsystems $(C_k,V_k^i,V_k^f,E_k,A_k,B_k,N)$ is stationary.

As before, we have an immediate corollary:

Corollary

A TPBVDS in BNF is displacement if and only if V_i and V_f are in the same block–diagonal form (6.4), (6.5) as E and A, and moreover, each of the subsystems $(C_k,V_k^i,V_k^f,E_k,A_k,B_k,N)$ is displacement.

Proof of Theorem 6.1

Consider a TPBVDS in BNF and assume without loss of generality that E and A are in block–form (6.1), (6.2), respectively. We first prove the following:

Lemma 6.1

The strong reachability and observability matrices of the overall system have the following form

$$R_s = \mathrm{diag}(R_s^1, \ldots R_s^M).W \tag{6.6}$$

$$O_s = Z.\mathrm{diag}(O_s^1, \ldots O_s^M) \tag{6.7}$$

where W and Z are invertible matrices and R_s^k and O_s^k are strong reachability and strong observability matrices of the k^{th} block of the system.

Proof

We begin by putting the pencil in standard form by premultiplying E and A by $(\alpha E+\beta A)^{-1}$ for some α and β. Note that $(\alpha E+\beta A)^{-1}$ is block–diagonal, as are the new E and A matrices. Indeed all we have done is to modify the system so that (6.3) is satisfied with all α_i equal to α and all β_i equal to β. Suppose $\alpha \neq 0$ (otherwise reverse the roles of E and A). It is not difficult to check in this case that the condition that no two blocks of E and A have the same eigenmode now implies that no two blocks of A have the same eigenvalue. Also in this case

$$\mathcal{R}_s = \mathrm{Im}[B|AB|\ldots|A^{n-1}B] \tag{6.8}$$

(replace E by $(I-\beta A)/\alpha$ in R_s and use the usual Cayley–Hamilton theorem). Equation (6.6) then follows from the usual causal system result. Equation (6.7) can be verified similarly.

Note that Lemma 6.1 demonstrates the equivalence of strong reachability/observability of the overall system and of all of the subsystems. Also, since every block is in standard form we can see that the strong reachability and observability spaces are E– and A–invariant, as when E and A are in standard form.

An examination of the proof of Theorem 2.1 shows that if we simply assume that E and A commute and that $\mathcal{R}_s$ and $\mathcal{O}_s$ are E– and A–invariant, the necessary and sufficient conditions for stationarity are

$$O_s[EV_iA-AV_iE]R_s = 0 \tag{6.9}$$

$$O_s[EV_fA-AV_fE]R_s = 0. \tag{6.10}$$

Consider next the following modification of our TPBVDS. Specifically, we keep C, E, A, B the same and simply null out the off-diagonal blocks of V_i and V_f. That is, let

$$V_i = \begin{bmatrix} V^i_{11} & \cdots & V^i_{1m} \\ \vdots & & \vdots \\ V^i_{M1} & \cdots & V^i_{MM} \end{bmatrix} \tag{6.11}$$

with the blocks of V_f defined similarly. Then let

$$\tilde{V}_i = \mathrm{diag}(V^i_{11}, V^i_{22} \ldots, V^i_{MM}) \tag{6.12}$$

$$\tilde{V}_f = \mathrm{diag}(V^f_{11}, V^f_{22} \ldots, V^f_{MM}). \tag{6.13}$$

What we wish to show is that $(C,\tilde{V}_i,\tilde{V}_f,E,A,B,N)$ is in BNF and has the same weighting pattern as the original system.

The fact that it is in BNF follows immediatly since we have not changed E and A and

$$\tilde{V}_iE^N+\tilde{V}_fA^N = V_iE^N+V_fA^N = I. \tag{6.14}$$

Thus what we need to show is that

$$O_s\tilde{V}_iR_s = O_sV_iR_s \tag{6.15}$$

$$O_s\tilde{V}_fR_s = O_sV_fR_s \tag{6.16}$$

or thanks to (6.6) and (6.7) that

$$O^k_sV^i_{kj}R^j_s = 0 \qquad j\neq k \tag{6.17}$$

$$O^k_sV^f_{kj}R^j_s = 0 \qquad j\neq k. \tag{6.18}$$

We focus on (6.17), as (6.18) follows similarly.

From (6.9) we immediately find that for $j\neq k$

$$O^k_s[E_kV^i_{kj}A_j]R^j_s = O^k_s[A_kV^i_{kj}E_j]R^j_s. \tag{6.19}$$

Recall that $\{E_j,A_j\}$ and $\{E_k,A_k\}$ are in standard form, and indeed by a block-diagonal transformation we can assume that $\alpha E_j+\beta A_j=\alpha E_k+\beta A_k=I$ for a single, given pair α and β. Without any loss of generality we can assume that this is true. Furthermore, assume that $\alpha\neq 0$ (otherwise reverse the roles of

E and A), so that

$$E_j = \gamma I+\delta A_j \quad , \quad E_k = \gamma I+\delta A_k. \tag{6.20}$$

Using (6.20) in (6.19) implies that

$$O_s^k[V_{kj}^i A_j]R_s^j = O_s^k[A_k V_{kj}^i]R_s^j. \tag{6.21}$$

Since $\mathcal{R}_s^j$ is A_j–invariant and $\mathcal{O}_s^k$ is A_k–invariant, we have that (6.21) implies that

$$O_s^k[V_{kj}^i p(A_j)]R_s^j = O_s^k[p(A_k)V_{kj}^i]R_s^j \tag{6.22}$$

for any polynomial p. Take any generalized eigenvector v of A_j in $\mathcal{R}_s^j$ corresponding to the eigenvalue λ_j of A_j. then there is an integer m so that

$$(\lambda_j I-A_j)^m v = 0. \tag{6.23}$$

Let $p(x)=(\lambda_j-x)^m$. Also, let w be any generalized left–eigenvector of A_k in $(\mathcal{O}_s^k)^\perp$ corresponding to the eigenvalue μ_k of A_k. Then, from (6.22) we have that

$$0 = w'V_{kj}^i p(A_j)v = w'p(A_k)V_{kj}^i v = (\lambda_j-\mu_k)^m w'V_{kj}^i v. \tag{6.24}$$

Thanks to (6.20) and the fact that $\{E_j,A_j\}$ and $\{E_k,A_k\}$ have no eigenmodes in common, $(\lambda_j-\mu_k)^m \neq 0$, so we can conclude that

$$w'V_{kj}^i v = 0. \tag{6.25}$$

But, since $\mathcal{R}_s^j$ is A_j–invariant and $\mathcal{O}_s^k$ is A_k–invariant, the columns of R_s^j and rows of O_s^k are spanned by such v's and w's, respectively, yielding (6.17).

Note that if the overall system (and therefore at least one of the subsystems) is not both strongly reachable and observable, there is some freedom in the choices of V_i and V_f. What the theorem says is that we can always choose these to be block–diagonal. If, however, all of the subsystems <u>are</u> strongly reachable and observable then the <u>only</u> possibility is for V_i and V_f to be block–diagonal. This is what happens in the Corollary (which, as before, corresponds to the case B=C=I). Note also, that since we can always take the boundary matrices to be block–diagonal, we have, as in the causal case, the fact that minimality of the overall system is equivalent to minimality of all of the subsystems.

There are several other important consequences of this theorem. First, note that for a stationary

TPBVDS in BNF with V_i and V_f as in (6.4), (6.5), Theorem 6.1 and Theorem 1, applied to each subsystem, allow us to deduce that in fact not only does (6.9) hold, but so does (2.13). This in turn allows us to obtain the simple form for the weighting pattern given in (2.40).

Lemma 6.1 allows us to study strong reachability and observability of individual eigenmodes. To see this, consider a TPBVDS transformed into the following normalized or block normalized form[1] where

$$E = \mathrm{diag}(E_1,\ldots,E_M) \tag{6.26a}$$

$$A = \mathrm{diag}(A_1,\ldots,A_M) \tag{6.26b}$$

where $\{E_i,A_i\}$ has a unique eigenmode σ_i with $\sigma_i\neq\sigma_j$ for $i\neq j$. Then we say that the eigenmode σ_j is strongly reachable if (E_j,A_j,B_j) is strongly reachable (i.e. R_s^j has full rank). It can easily be verified that σ_j is strongly reachable if and only if

$$[\sigma_j E-A|B]$$

has full row rank ($\sigma_j=\infty$ is strongly reachable if and only if $[E|B]$ has full row rank). Similarly, we say that an eigenmode σ_j is strongly observable if (C_j,E_j,A_j) is strongly observable (i.e. O_s^j has full rank). Eigenmode σ_j is strongly observable if and only if

$$\begin{bmatrix}\sigma_j E-A\\ C\end{bmatrix}$$

has full column rank ($\sigma_j=\infty$ is strongly observable if and only if $\begin{bmatrix}E\\ C\end{bmatrix}$ has full column rank). In the displacement case, the boundary matrices are also in block diagonal form:

$$V_i = \mathrm{diag}(V_1^i,\ldots,V_M^i) \tag{6.27a}$$

$$V_f = \mathrm{diag}(V_1^f,\ldots,V_M^f). \tag{6.27b}$$

The BNF (6.26)–(6.27) allows us to consider weak reachability and observability of individual eigenmodes. An eigenmode σ_j is called weakly reachable (observable) if subsystem j is weakly

---

[1]We can always transform any regular $\{E,A\}$ into the block form (6.26). Assume $\{E,A\}$ is in standard–form, then we find T such that TAT^{-1} and TET^{-1} are in real Jordan form (thanks to the standard–form, E and A can be put into Jordan form simultaneously). Then (6.26) can be obtained by reordering the Jordan blocks of TAT^{-1} and TET^{-1}.

reachable (observable). Also σ_j is weakly reachable if and only if

$$[\sigma_j E - A | V_i B | V_f B]$$

has full row rank; it is weakly observable if and only if

$$\begin{bmatrix} \sigma_j E - A \\ CV_i \\ CV_f \end{bmatrix}$$

has full column rank.

We can also use Theorem 6.1 to obtain the following result:

Theorem 6.2

Consider a minimal, stationary TPBVDS, then any eigenmode of the strongly unreachable (unobservable) part of the system is also an eigenmode of the strongly reachable (observable) part of the system.

Proof

Suppose that σ_k is an eigenmode of the strongly unreachable part of the system but not of the strongly reachable part. Theorem 6.1 allows us to break–down the system into subsystems each one of which has a distinct eigenmode. In particular, let $\Sigma_k = (C_k, V_k^i, V_k^f, E_k, A_k, B_k, N)$ denote the subsystem associated to eigenmode σ_k. Then, since Σ_k is minimal, it has a strongly reachable part (otherwise, B_k must be zero, the subsystem has weighting pattern 0 and the minimal realization has dimension 0). Thus, σ_k is an eigenmode of the strongly reachable part of Σ_k and of the original system. This, of course, is a contradiction.

Before closing this section, we should mention that the motivation behind introducing the concepts of block standard form and BNF has been the usefulness of the following block standard form

$$E = \mathrm{diag}(I, I, A_b) \tag{6.28a}$$

$$A = \mathrm{diag}(A_f, U, I) \tag{6.28b}$$

where the eigenvalues of A_f and A_b are all inside the unit circle and the eigenvalues of U on the unit circle. This particular block standard form has been used in [2,3,16] and shall be used in the next chapter for studying the stability and the stochastic realization problem for TPBVDS's. When $\{E,A\}$ has no eigenmode on the unit circle, the block standard form (6.28) is called the forward–backward stable form.

7–Conclusions

In this report we have developed some of the system–theoretic properties of two–point boundary–value descriptor systems. We have derived detailed characterizations of reachability, observability, and minimality with particular attention paid to the shift–invariant case. As had already been noted for continuous–time, non–descriptor boundary–value systems, minimality for TPBVDS's is a bit more complicated than for causal systems. Indeed there is a certain degree of nonuniqueness in minimal realizations. One open problem that we have noted concerns whether one can use this freedom to guarantee that a displacement system always has a minimal realization that is also displacement.

Another concept that we have introduced and studied in this paper is extendibility, i.e. the idea of thinking of a TPBVDS as being defined on a sequence of intervals of increasing length. Once one introduces such a notion, it becomes possible to talk about the realization problem [28] and asymptotic properties such as stability [16].

References

1. R. Nikoukhah, A.S. Willsky, and B.C. Levy, "Boundary–value descriptor systems: well–posedness, reachability and observability," Int. J. Control, vol. 46, no. 5, pp. 1715–1737, Nov. 1987.

2. R. Nikoukhah, "System Theory for two–point boudary value descriptor systems," M.S. thesis, Dept. of Elec. Eng. and Comp. Science, and Report LIDS–TH–1559, Lab. for Information and Decision Systems, M.I.T., Cambridge, MA, June 1986.

3. R. Nikoukhah, M.B. Adams, A.S. Willsky, and B.C. Levy, "Estimation for Boundary–value descriptor systems," Report LIDS–P–1600, Lab. for Information and Decision Systems, M.I.T., Cambridge, MA, Aug. 1986; to appear in Circuits, Systems, and Signal Processing.

4. D.G. Luenberger, "Dynamic systems in descriptor form," IEEE Trans. on Automat. Contr., vol. AC–22, no. 3, pp.312–321, June 1977.

5. D.G. Luenberger, "Time–invariant descriptor systems," Automatica, vol. 14, no. 5, pp. 473–480, Sept. 1978.

6. A.J. Krener, "Boundary value linear systems," Asterisque, vol. 75–76, pp. 149–165, Oct. 1980.

7. A.J. Krener, "Acausal realization theory, Part I: Linear deterministic systems," SIAM J. Control and Optimization, vol. 25, no. 3, pp. 499–525, May 1987.

8. M.B. Adams, "Linear estimation of boundary value stochastic processses," Sc.D. Thesis, Dept. of Aero. and Astro. MIT, Cambridge, MA., April 1983.

9. M.B. Adams, A,S Willsky, and B.C. Levy, "Linear estimation of boundary value stochastic processes, Part 1: The role and construction of complementary models," IEEE Trans. Automat. Contr., vol. AC–29, no. 9, pp. 803–811, Sept. 1984.

10. M.B. Adams, A,S Willsky, and B.C. Levy, "Linear estimation of boundary value stochastic processes, Part 2: 1–D smoothing problems," IEEE Trans. Automat. Contr., vol. AC–29, no. 9, pp. 803–811, Sept. 1984.

11 I. Gohberg and M.A. Kaashoek, "On minimality and stable minimality of time varying linear systems with well posed boundary conditions," Int. J. Control, vol. 43, no. 5, pp. 1401–1411, 1986.

12 I. Gohberg and M.A. Kaashoek, and L. Lerer, "Minimality and irreducibility of time–invariant linear boundary–value systems," Int. J. Control, vol. 44, no. 2, pp. 363–379, 1986.

13 I. Gohberg and M.A. Kaashoek, "Similarity and reduction for time–varying linear systems with well–posed boundary conditions," SIAM J. Control and Optimization, vol. 24, no. 5, pp. 961–978, Sept. 1986.

14 W.M. Wonham, Linear Mutivariable Control: A Geometric Approach, Springer–Verlag, New York, 1979.

15. I. Gohberg, M.A. Kaashoek, and L. Lerer, "On minimality in the partial realization problem," Systems and Control Letters, vol. 9, pp. 97–104, 1987.

16. R. Nikoukhah, B.C. Levy, and A.S. Willsky, "Stability, stochastic stationarity, and generalized Lyapunov equations for two–point boundary–value descriptor systems," submitted for publication.

17. F. L. Lewis, "Adjoint matrix, Cayley–Hamilton theorem, and Fadeev's method for the matrix pencil (sE–A)," Proc. 22nd IEEE Conf. Dec. and Contr., San Antonio, TX, 1983.

18. F.L. Lewis, "Descriptor systems: expanded descriptor equations and markov parameters," IEEE Trans. on Automat. Contr., vol. AC–28, no.5, May 1983.

19. F. L. Lewis, "Descriptor systems: decomposition into forward and backward subsystems," IEEE Trans. on Automat. Contr., vol. AC–29, no.1, Jan. 1984.

20. F. L. Lewis, "Fundamental reachability, and observability matrices for discrete descriptor systems," IEEE Trans. on Automat. Contr., vol. AC–30, no.5, May 1985.

21. G. C. Verghese, B. C. Levy, and T. Kailath, "A generalized state–space for singular systems," IEEE Trans. on Automat. Contr., vol. AC–26, no.4, August 1981.

22. P. M. Van Dooren, "The generalized eigenstructure problem in linear system theory," IEEE Trans. on Automat. Contr., vol. AC–26, no.1, Feb. 1981.

23. E. L. Yip and F. Sincovec, "Solvability, controllability, and observability of continuous descriptor systems," IEEE Trans. on Automat. Contr., vol. AC–26, no.3, June 1981.

24. G. C. Verghese, Infinite Frequency Behaviour in Generalized Dynamincal Systems, Ph.D. Thesis, Stanford University, Dec. 1978.

25. H. H. Rosenbrock, "Structural properties of linear dynamical systems," Int. J. Contr., vol. 20, no.2, 1974.

26. F. R. Grantmacher, The Theory of Matrices. New York: Chelsea, 1974.

27. T. Kailath, Linear Systems. Prentice–Hall, 1980.

28. R. Nikoukhah, A. S. Willsky and B. C. Levy, "Realization theory for deterministic two–point boundary–value descriptor systems," in preparation.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique